Damaged Book

Pages Miis Within The Book Only

CHAPTER I

MAGNETISM (INTRODUCTORY)

Magnets and Magnetic Poles.—The original observations which led to the development of the study of Magnetism are—the setting in one particular direction of a piece of lodestone or magnetite when freely suspended, and the picking up of small pieces of magnetite or iron by a larger piece of magnetite. Further, a similar polarity is produced in pieces of steel by the simple process of rubbing from one end to the other with that part of the magnetite where iron filings cling in greatest quantity. The piece of steel will then set with one particular end pointing approximately north when suspended so that it is free to turn in a horizontal plane, and it will pick up iron filings.

A knitting needle may be used for the purpose of the above experiment, and by magnetising several such needles and suspending each of them in turn in a stirrup supported by a single silk fibre, the ends which point to the north may be determined. These ends are called the north-seeking poles, or, more shortly, the N poles of the needles; the other ends are of course the south-seeking or S poles. If then one of the needles be suspended and the poles of one of the others brought in turn near its poles, it will be seen that two N poles repel each other, as do two S poles, but that a N pole and a S pole attract each other.

When two magnetic poles are brought near to each other, the force between them is the most important guide we have with regard to the strength of the poles; in fact, we can only say that two poles are equal when they experience equal forces on being brought in turn into identical positions with respect to a third pole. If we follow out this conception and imagine a number of equal poles to be produced, we shall then see that by combining these arbitrary unit poles, which we will imagine to be all of one kind, to form two poles A and B, the force between A and B is proportional to the product of the number of units in each; for if either be increased n times by the addition of more units, the force is also increased n times, provided that the force between any two units is in no way affected by the presence of other poles. Experience tells us that magnetic forces are in this respect like gravitational forces; the ... third body does ...

two. Hence we may say that the force between two magnetic poles varies as the product of th[illegible]ngths.

The Inverse Square Law.—T[illegible]w of variation with their distance apart, of the force betw[illegible] two poles must be determined experimentally. But here again a [illegible] f the case of gravity helps us. If we imagine t[illegible]es each concentrated at a point, the force bet[illegible] will vary inversely as the square of their distanc[illegible]. The experimental proof was first undertaken by Coulomb, who, using his torsion balance, showed the law to be true within about three parts per hundred.

A similar degree of accuracy may be obtained with the Hibbert magnetic balance, but all these direct and simple proofs are made on the assumption that the poles of magnets are situated at or near definite points close to the ends of the magnets, whereas this is never the case. In the experiment of picking up the iron filings or of approaching one magnet to another to observe the force between poles, it is evident that the magnetic effects extend over large parts of the surfaces of the magnets, being very small or zero near the middle and increasing towards the ends. The conception of point poles, however, is a very important one, and we have every reason to believe that for such poles the law of force is the inverse square law; that is

$$\text{Force} = A\frac{m_1 m_2}{r^2}$$

where m_1 and m_2 denote the strengths of the poles measured in any arbitrary units and r is their distance apart. If the N pole be given a positive sign and the S pole a negative one, it follows if A is positive that a positive force is a repulsion, whereas an attraction is negative. The direct experimental proof of this law is impossible, but we shall see on [illegible] that the experiments of Gauss establish it with a fair degre[illegible] accuracy. The most important reason for accepting the truth of the law lies in the fact that, *without exception*, effects calculated on the assumption of its truth are in accordance with experimental results, always within the limits of accuracy of which the experimental work is capable.

Units.—In choosing our units, those of *force* and *distance* are already fixed for us on the scientific system, otherwise known as the Centimetre-Gramme-Second system. It is therefore most convenient to choose our *unit of magnetic pole* so that the constant A in our equation becomes unity, and this will be the case if th[illegible] unit values of m_1 and m_2 are such that the force between t[illegible] poles is one dyne when their distance apart is one cent[illegible]. The medium in wh[illegible] [illegible] are situat[illegible] [illegible]

what is nearly the same thing for magnetic purposes—in air. Thus the unit pole may be defined as the pole which placed in air one centimetre from a · equal pole repels or attracts it with a force of one dyne ; and - the force between any two poles whose strengths are meas terms of these units,

$$\text{Force} = \frac{m_1 m_2}{r^2} \text{ dynes} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Magne ·lation (1) by itself is of very little use ; we can only find ns of it, the force on a pole when the magnitude and position ll other poles are known. These are never known in any real case ; in fact, the force may not, strictly speaking, be due to " poles " at all ; and yet it is very important to be able to express the force on the given pole in terms of external effects. The resultant of all forces acting on the pole for any given arrangement of magnetic bodies depends upon its position, and if the pole be a N pole of unit strength, the force upon it is called the *Strength of Magnetic Field* at the point, or the *Magnetic Force*, or *Intensity*, the symbol usually used to denote it being H.

It follows that the force on any pole of strength m is equal to Hm dynes, or

$$F = Hm \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

It must be noticed that H is a vector quantity like a force, and is subject to the law of addition of vectors, sometimes known as the law of the parallelogram of forces.

We can now calculate the magnetic field in several simple cases, the general process being to imagine a unit N pole to be placed at the point at which we require the field, calculate the force upon it due to each known pole, and then find the resultant.

Thus the field at a distance r cm. from a N pole of strength m is $+\frac{m}{r^2}$; for putting $m_1 = +m$, and $m_2 = +1$ in equation (1) we have $F = \frac{m}{r^2}$, the strength of field required.

Field due to Magn Case (i).

Let the point P (Fig. 1), at which the field is required, be situated on the line joining the poles of the magnet and at a distance d from its middle point.

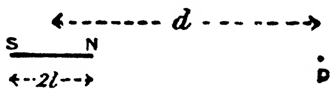


Fig. 1.

If m be the strength of pole of the magnet and l half its length,

$$\text{Field at P due to N} = \frac{m}{(d-l)^2}$$

$$\text{,, \quad ,, \quad ,, \quad S} = -\frac{m}{(d+l)^2}.$$

Since these two are in the same line,

$$\text{Resultant field} = \frac{m}{(d-l)^2} - \frac{m}{(d+l)^2} = \frac{4mld}{(d^2-l^2)^2}.$$

If the length of the magnet is so small that l^2 is negligible in comparison with d^2,

$$\text{Resultant field} = \frac{4ml}{d^3}.$$

Case (ii).

P being situated on a line bisecting the magnet at right angles (Fig. 2), the field PA due to N has strength $\frac{m}{d^2+l^2}$, since PN$=\sqrt{d^2+l^2}$.

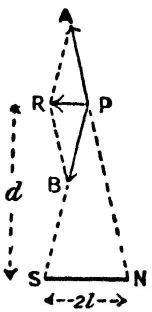


Fig. 2.

Similarly field due to S, $\text{PB} = \frac{m}{d^2+l^2}$.

The resultant field is evidently PR, and from the geometry of the figure we see that

$$\frac{\text{PR}}{\text{PA}} = \frac{\text{NS}}{\text{PN}}$$

or,

$$\text{PR} = \text{PA} \cdot \frac{\text{NS}}{\text{PN}}$$

thus, resultant field

$$= \frac{m}{d^2+l^2} \cdot \frac{2l}{\sqrt{d^2+l^2}}$$

$$= \frac{2ml}{(d^2+l^2)^{\frac{3}{2}}}$$

As in Case (i), if l^2 is small in comparison with d^2,

$$\text{resultant field} = \frac{2ml}{d^3}.$$

Magnetic Moment.—We are now in a position to deal with the case of a suspended magnet, free to turn in a horizontal plane. Such a magnet comes to rest with its N pole pointing north. The magnet is evidently situated in a magnetic field, known in this case as the Earth's Field; and assuming this to be equivalent to a horizontal magnetic field of strength H, the N pole of the suspended magnet experiences a force $+\text{H}m$ and the S pole a force $-\text{H}m$, the two giving rise to a couple whose turning moment is equal to either force multiplied by the perpendicular distance AN between them (Fig. 3).

$$\therefore \text{Couple} = \text{H}m(\text{AN})$$
$$= \text{H}m \cdot \text{NS} \cdot \sin\theta.$$

This couple vanishes when the magnet has a position parallel to H, the magnetic meridian, that is, when $\theta=0$; which explains the setting of the compass needle in a N and S direction.

The expression for the couple acting on the needle consists of three parts, H depending on the Earth, θ on the position of the magnet, and m . NS depending on the magnet itself. The last quantity is called the magnetic moment M of the magnet. In the case of a fictitious magnet consisting of two point poles, it is the *strength of pole multiplied by the distance between the poles*, but the definition of magnetic moment from the expression

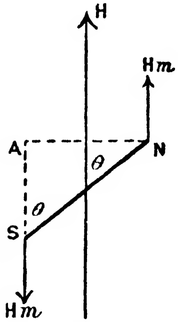


FIG. 3.

$$\text{Couple}=\text{HM}\sin\theta \quad . \quad . \quad . \quad (3)$$

does not depend upon any such fiction, for any magnet may be suspended in a magnetic field and the couple required to maintain it in a given position measured, and M therefore found.

If the magnet be maintained at right angles to the magnetic field, $\theta=90°$ and $\sin\theta=1$.

$$\therefore \text{Couple}=\text{HM},$$

and we may, from this, define the moment of a magnet as *the couple required to maintain it at right angles to a magnetic field of unit strength.*

We see that the expressions for the field due to a magnet become

$$\text{Case (i)} \quad \frac{2\text{M}d}{(d^2-l^2)^2} \quad \text{or} \quad \frac{2\text{M}}{d^3}$$

$$\text{Case (ii)} \quad \frac{\text{M}}{(d^2+l^2)^{\frac{3}{2}}} \quad \text{or} \quad \frac{\text{M}}{d^3}$$

where M is substituted for $2ml$, the length of the magnet being $2l$. An ordinary magnet cannot be said to have any definite length, as the pole is a collection of point poles, but if situated in a uniform field, the centre of force for all the parallel forces may be found, just as in a case of finding the centre of gravity of a body, and the distance between these two effective poles multiplied by the strength of either, may be seen to lead to the same definition of magnetic moment as was derived from the consideration of the couple in uniform field.

The Magnetometer.—The position of equilibrium of a magnetised needle suspended in two magnetic fields at right angles to each other may now be found.

H and F being the strengths of the respective fields and M the

magnetic moment of the magnet, MH $\sin\theta$ is the couple tending to rotate it into the direction of H, and MF $\sin(90°-\theta)$ = MF $\cos\theta$ is the couple tending to rotate it in the direction of F (Fig. 4). The needle is therefore in equilibrium when these couples are equal, *i.e.* when

$$MH\sin\theta = MF\cos\theta$$

or,

$$\frac{F}{H} = \tan\theta.$$

The same result might have been obtained by remembering that the magnet will set in the direction of the resultant field, and that the resultant is inclined at an angle $\tan^{-1}\frac{F}{H}$ to the field H.

The field F may be due to a variety of causes; later, when considering galvanometers we shall have to treat it as due to an electric current in a coil of wire, but in the present case we may consider it to be due to a bar magnet.

The magnet being situated E or W of, and at a distance d from,

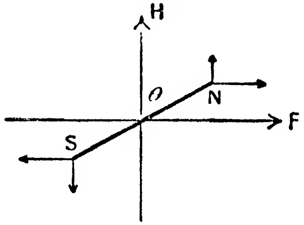


Fig. 4.

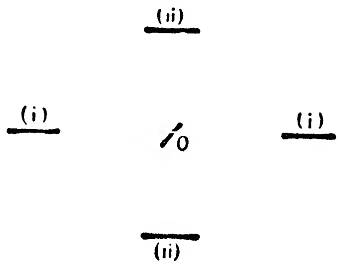


Fig. 5.

the suspended needle O (Fig. 5, i), the field at O, due to the magnet, is $\frac{2M}{d^3}$, and hence the needle O will come to rest when at an angle θ to its position of equilibrium with the magnet absent.

Then,

$$\frac{2M}{d^3H} = \tan\theta,$$

or,

$$\frac{M}{H} = \frac{d^3}{2}\tan\theta.$$

If the position (ii) be employed the magnet is situated N or S of the needle but still pointing E and W.

Then,

$$\frac{M}{d^3H} = \tan\theta,$$

or.

$$\frac{M}{H} = d^3\tan\theta.$$

If the length of the magnet is not so small that its square may be neglected in comparison with the square of the distance between the magnet and the needle, the more exact formulæ must be used, *i.e.*,

Case (i) $$\frac{M}{H}=\frac{(d^2-l^2)^2}{2d}\tan\theta,$$

Case (ii) $$\frac{M}{H}=(d^2+l^2)^{\frac{3}{2}}\tan\theta.$$

Fig. 6 shows a common form of simple magnetometer for carrying out the measurements of deflection. The needle is a short one and is attached to a light pointer, which may be a fine piece of aluminium wire, or a piece of glass tubing drawn out fine while soft. The suspension may be a silk fibre, or a needle-point bearing in an agate cup. The whole instrument is placed so that the ends of the pointer are at 0°—0° on the scale when no deflecting magnet is present.

There are several sources of error, but their effects may be

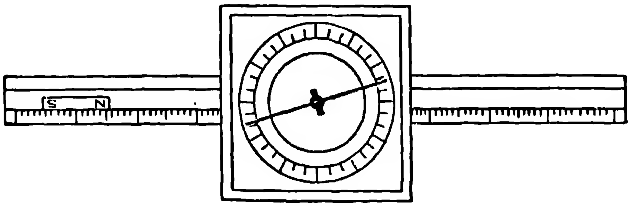


FIG. 6.

eliminated by taking a series of readings, provided that the errors themselves are small.

(*a*) The point of suspension may not be at the centre of the circular scale, and therefore both ends of the pointer are read.

(*b*) The *deflecting* magnet may not be symmetrically magnetised. To eliminate this error the magnet is turned over so that its N and S poles change places and the readings are taken again.

(*c*) The point of suspension may not be at the zero of the long straight scale, and therefore the magnet must now be placed at an equal distance, according to the scale, on the other side of the needle and the previous readings repeated.

In this way eight readings are made and the mean is free from the errors mentioned. In making the instrument, care must be taken that the pointer is at right angles to the needle. If this is not done the zero line when the instrument is set up will not be at right angles to the meridian, and the magnet will not be in such a position that the fields due to earth and magnet are at right angles. Consequently our equations do not apply. In order to make sure that the line joining the ends of the pointer

is at right angles to the magnetic axis of the needle, the needle should be suspended and the positions of the ends of the pointer marked. Then the system should be turned over and suspended from the other side. If now the pointer covers its first position it must be at right angles to the magnetic axis, but if it does not, it must be bent or in some way moved until it indicates the same reading whichever way up the needle is suspended.

By taking various distances and using Cases (i) and (ii) in turn, the relations $\frac{M}{H}=\frac{d^3}{2}\tan\theta$, and $\frac{M}{H}=d^3\tan\theta$, or, if the magnet is not short enough, the more exact relations, may all be verified.

Again, since the quantity $\frac{M}{H}$ has been found, we may, by changing the magnet for another, find the ratio of the two magnetic moments.

$$\frac{M_1}{H}=d_1^3\tan\theta_1, \qquad \frac{M_2}{H}=d_2^3\tan\theta_2$$

$$\therefore \frac{M_1}{M_2}=\frac{d_1^3\tan\theta_1}{d_2^3\tan\theta_2}$$

or, by using the same magnet and transferring the magnetometer from one place to another, we may find the ratio of the earth's horizontal magnetic fields at the two places.

$$\frac{M}{H_1}=d_1^3\tan\theta_1 \qquad \frac{M}{H_2}=d_2^3\tan\theta_2$$

$$\therefore \frac{H_1}{H_2}=\frac{d_2^3\tan\theta_2}{d_1^3\tan\theta_1}.$$

The form of the moving part in the simple magnetometer does not allow of great accuracy in observing the deflection, for the thickness of the pointer itself is quite a large fraction of the size of a division of an ordinary scale of degrees. Although error due to parallax is avoided by fixing the scale on a piece of plane mirror so that the eye may always be kept vertically over the scale, the image of the pointer in the mirror and the pointer itself being made to coincide, there is still the fact that the thickness of the pointer is perhaps $\frac{1}{20}$ of the total deflection to be read. To make the scale larger would mean using a longer pointer, and thus a larger apparatus; the increase in weight of the pointer would require a stouter support, which again would mean a loss of sensitiveness. What is wanted is a long weightless pointer, and fortunately this is exactly what we have in a beam of light. As the method of a reflected beam is so largely employed in the

case of galvanometers as well as for magnetometers we will consider it somewhat in detail.

Fig. 7 is a plan of the arrangement. F is the filament of some form of electric glow lamp, and its position is near the principal focus of the lens L, which is merely a condensing lens, to bring the light from the filament into a suitable direction. The magnetometer mirror is usually concave, having a radius of curvature of about a metre, and produces upon the scale S an image of a vertical scratch upon the lens. The magnetic "needle" consists of a few pieces of magnetised watch spring attached to the back of the mirror as shown at M′. The distance LS upon the scale, where L is supposed to be the middle of the scale, is a measure of the deflection, and for many purposes this is all that we require. But if the actual angular deflection is required, the distance LM from the mirror to the scale must be found; and remembering that the reflection occurring at the mirror doubles

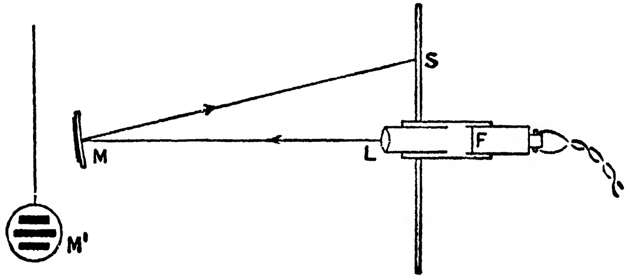


FIG. 7.

the rotation of the beam of light, the actual rotation of the mirror is $\frac{1}{2}\tan^{-1}\frac{\text{SL}}{\text{LM}}$. In most cases the deflection is so small that the angle and its tangent do not differ greatly, and then we may take the deflection as proportional to SL. Sometimes, instead of the lamp and lens we have a telescope, and in this case (Fig. 8) the suspended mirror is plane instead of concave, the telescope being focused upon the image of the scale in the mirror. The position of the cross wire in the eye-piece as seen upon the image of the scale, enables us to observe the deflection of the needle.

Gauss's Proof of the Inverse Square Law.—The two expressions for the strength of field near a magnet, $\frac{2\text{M}d}{(d^2-l^2)^2}$ for Case (i) and $\frac{\text{M}}{(d^2+l^2)^{\frac{3}{2}}}$ for Case (ii), are both obtained on the assumption of the inverse square law, and the resulting equations for the

magnetometer, $\frac{M}{H}=\frac{(d^2-l^2)^2}{2d}\tan\theta$ and $\frac{M}{H}=(d^2+l^2)^{\frac{3}{2}}\tan\theta$, also in their turn depend upon the truth of the law.

In either case, if we measure θ for different values of d, we may prove the constancy of $\frac{M}{H}$, and thus demonstrate the truth of the inverse square law. It should be noticed that l the half-length of the magnet is not accurately known, since the poles are not at the ends, neither are they point poles. However, l may be found to a first approximation by taking two readings for θ and d, in one case, and equating the values of $\frac{M}{H}$ obtained. We thus have an equation in l, and this may then be calculated and substituted in the other determinations. This method is only an approximation; it is better to use an exceedingly sensitive

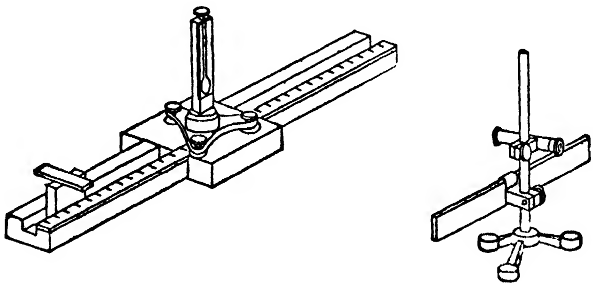


FIG. 8.

magnetometer so that reasonably accurate readings of the deflection may be made, with a magnet so small and so far distant that the relations $\frac{M}{H}=\frac{d^3}{2}\tan\theta$, and $\frac{M}{H}=d^3\tan\theta$, may be used. With a length of magnet of 4 cm. at a distance of 50 cm. from the needle, $d^2=2500$ and $l^2=4$, and thus the error involved in neglecting l^2 in comparison with d^2 is about 0·16 per cent. In practice the deflection might amount to, say, 15 scale divisions, with a probable error of one-tenth of a division. It is thus seen that the error introduced by neglecting l^2 is decidedly less than the unavoidable error in reading the deflection.

Employing the position of Case (i), Gauss observed the deflection for a given magnet at a given distance. He called this the "A" position. Next placing the magnet in the position of Case (ii), which he called the "B" position, the deflection was again observed.

Since, $\frac{M}{H}=\frac{d^3}{2}\tan\theta_1 \ldots$ (A), and $\frac{M}{H}=d^3\tan\theta_2 \ldots$ (B)

it follows that
$$\frac{\tan\theta_1}{\tan\theta_2}=2$$

if the equations are correct. If the law of attraction were an inverse law of any other power than 2, let us say n, it may then be shown that

$$\frac{\tan\theta_1}{\tan\theta_2}=n.$$

For, referring to Fig. 1,

$$\text{Field at P due to N}=\frac{m}{(d-l)^n}$$

$$\text{Field at P due to S}=\frac{m}{(d+l)^n}$$

$\therefore$ resultant field

$$=m\frac{(d+l)^n-(d-l)^n}{(d^2-l^2)^n}$$

$$=md^n\frac{\left(1+\frac{l}{d}\right)^n-\left(1-\frac{l}{d}\right)^n}{(d^2-l^2)^n}$$

$$=md^n\frac{\left\{1+\frac{nl}{d}+\frac{n(n-1)}{1.2}\cdot\frac{l^2}{d^2}+\ldots-1+\frac{nl}{d}-\frac{n(n-1)}{1.2}\cdot\frac{l^2}{d^2}+\ldots\right\}}{(d^2-l^2)^n}$$

$$=2md^n\frac{\left\{\frac{nl}{d}+\frac{n(n-1)(n-2)}{1.2.3}\cdot\frac{l^3}{d^3}+\ldots\right\}}{(d^2-l^2)^n}$$

Now if l^2 is negligible in comparison with d^2, then $\frac{l^3}{d^3}$ and higher powers of $\frac{l}{d}$ are negligible in comparison with $\frac{l}{d}$, and the expression for the resultant field simplifies to

$$\frac{2mln}{d^{n+1}}=\frac{nM}{d^{n+1}}$$

For the "B" position of Gauss, referring to Fig. 2—

$$PA=\frac{m}{(d^2+l^2)^{\frac{n}{2}}}$$

$$\therefore \text{ resultant field}=\frac{m}{(d^2+l^2)^{\frac{n}{2}}}\cdot\frac{2l}{(d^2+l^2)^{\frac{1}{2}}}$$

$$=\frac{2ml}{(d^2+l^2)^{\frac{n+1}{2}}}$$

If now l^2 in the denominator be neglected—

$$\text{Field} = \frac{2ml}{d^{n+1}} = \frac{M}{d^{n+1}}$$

and it follows that the deflections in the " A " and " B " positions of Gauss should be so related that—

$$\frac{\tan \theta_1}{\tan \theta_2} = n.$$

In the original paper of Gauss [1] the couple is calculated for any relative positions of the two magnets, and for the purpose of the experiment is reduced to the simple forms of the " A " and " B " positions. d varies from 1·1 metre to 4·0 metres and the deflection from 1° 57′ 24·8″ to 0° 2′ 22·2″ and the values calculated on the assumption of the law $F \propto \frac{m_1 m_2}{d^2}$ agree with the observed results to within a few seconds, thus proving that the force varies as the product of the pole strengths and inversely as the second power of their distance apart.

Lines and Tubes of Force.—A field of force such as a magnetic field, a gravitational field, or an electric field may, as we have seen, be completely defined at every point in terms of the force which would be exerted upon unit quantity of magnetic pole, matter, or, as we shall see later, electricity, if placed at each point in turn. If we imagine a free N pole placed at any point in a magnetic field, it will experience a force in the direction of the field, and on allowing it to move freely, it will evidently follow a path whose direction is, at each point, the direction of the field. Such a path is called a *line of force*. The conception of a line of force is important, as it naturally leads us to look to the medium in which the poles are situated, for the explanation of the forces between them, and it is this fact which makes the work of Faraday of such enormous importance. If at each point of the field, lines of force be drawn so that the number of lines per square centimetre is numerically equal to the strength of field at the point, and the process continued, it may be shown that such lines are continuous curves, since they satisfy the same condition as the stream lines in the space occupied by a moving liquid ; in fact, their resemblance to such lines is a very close one. Owing to the discontinuity of such lines in space, it is sometimes preferred to surround each line by a tube, such that the tubes touch each other laterally and fill the whole of space. Thus the lines or tubes of force, by their direction, indicate the direction of the field, and by the closeness with which they are packed (number per square centimetre) the strength of the field. These tubes

[1] C. F. Gauss, *Poggend. Ann.*, **38**, p. 591. 1833.

are not identical with the Faraday tubes of force, which will be described in Chapter IV.

If the tubes or lines tend to shrink in length, and at the same time to expand laterally, just as tubes of a solid material under tensile strain would do, the forces between poles would follow; but we must be careful not to push the analogy too far. Although the idea may be a useful one in concentrating our attention upon the medium rather than the poles, we must keep quite an open mind as to the nature of the medium.

The tension in the lines would tend to pull N and S together (Fig. 9, i), while the lateral push of the lines, together with the pull of the lines to each side, would also tend to urge N and N apart (Fig. 9, ii).

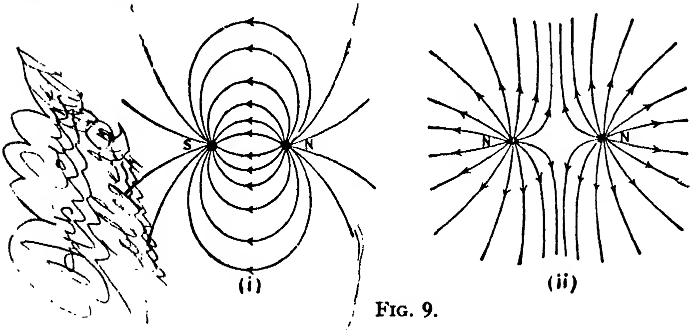


Fig. 9.

Potential.—There is another way of defining a field of force. Just as the flow of heat occurs in the direction of greatest variation in temperature, so the resultant direction of the magnetic field is that in which a quantity which we shall call *magnetic potential* varies most rapidly. *Potential may be defined as a quantity whose space rate of variation in any direction is the strength of the field in that direction.* This idea is common to all fields of force, and most of the results obtained here may be transferred, with mere alteration of the names of the quantities, to problems in gravitation, electricity, etc.

From our definition of potential we see that if V is the potential at any point x, $\frac{dV}{dx}$ is the rate of change of potential as we pass from point to point, or the ratio of difference of potential to distance travelled, for a very small path. If then by definition of V, this quantity is the strength of field at the place considered, $F=-\frac{dV}{dx}$, the use of the negative sign being conventional and indicating that in magnetic problems the force between like poles is a repulsion, the potential diminishing as the distance from the pole increases.

If a unit pole be placed at x, Fig. 10, the force experienced by it is given by the above expression, $F=-\frac{dV}{dx}$. The work done for a small movement dx is then

FIG. 10.

$$Fdx=-\frac{dV}{dx}dx=-dV,$$

$$\therefore V_B^A=V_A-V_B=-\int_B^A Fdx$$

which means that the difference in the potential between the two points A and B is the work done in carrying a unit magnetic pole from one point to the other.

Now consider the force to be due to a N pole of strength m situated at N. The strength of field at x due to this is $\frac{m}{x^2}$, or

$F=\frac{m}{x^2}$,

$$V_B^A=-\int_B^A \frac{m}{x^2}dx=\left[\frac{m}{x}\right]_B^A=\frac{m}{A}-\frac{m}{B}.$$

Potential can therefore only be measured by its differences, as there is no absolute zero of potential, and consequently we cannot speak of the absolute potential of any point. Nevertheless it should be noted that there is no difference of potential between two points, that is, they are at the same potential, when a magnetic pole may be conveyed from one of the points to the other without the expenditure of work. At an infinite distance from all poles, the forces are zero, and therefore all points are at the same potential. If we choose as the zero from which potential shall be measured, this potential at infinity, we see on putting $B=\infty$ in our equation, that $V=\frac{m}{A}$, where V is now the potential at A due to the pole m.

The potential at a point may therefore be defined as the work done in bringing a unit N pole from infinity to the point.

The potential at distance r from a N pole of strength m is $\frac{m}{r}$, and at distance r from an equal S pole it is $-\frac{m}{r}$.

Further, the work done in carrying a unit pole from any one point to any other is independent of the path by which the pole is taken: for if BCA be any path from B to A (Fig. 11), the work done is the same as in travelling from B to A′ along the straight line BDA′, provided that A and A′ are equidistant from N. The path A′A, which is an arc of a circle, is everywhere at

right angles to the field due to N, so that no work is done in carrying the unit pole from A′ to A. Hence the work for all paths such as BCA is the same as that for the path BDA′, and is therefore constant.

The same conclusion is reached if we imagine the unit pole to be carried round any closed path, such as ACBDA′A. The total work done is zero, for everything is now in the same condition as at the start, and therefore the work done for the path ACB is equal and opposite to that for the path BDA′A, and is therefore equal to that for the path AA′DB. Hence whatever path we take from A to B the work done is the same in amount.

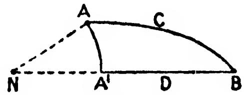


FIG. 11.

Potential due to Magnet.—Referring to Fig. 1 we see that the potential at P due to N is $+\frac{m}{d-l}$; and potential at P due to S is $-\frac{m}{d+l}$.

$$\therefore \text{ Actual potential at P} = \frac{m}{d-l} - \frac{m}{d+l}$$
$$= \frac{2ml}{d^2-l^2}$$
$$= \frac{M}{d^2-l^2}$$

For a very short magnet the potential becomes $\frac{M}{d^2}$.

In Fig. 2

$$\text{Potential at P due to N} = +\frac{m}{\sqrt{d^2+l^2}}$$
$$\text{,, \quad ,, \quad ,, \quad S} = -\frac{m}{\sqrt{d^2+l^2}}.$$

As these values are equal and opposite it follows that every point on the line bisecting the magnet at right angles is at zero potential. For a point P on a line passing through the middle of the magnet and inclined at an angle θ to the magnet, drop perpendiculars from N and S to OP (Fig. 12).

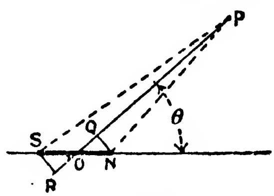


FIG. 12.

Then if the magnet is very small compared with the distance OP, we may without sensible error write

$$QP = NP, \text{ and. } RP = SP.$$

Then, Potential at $P=\frac{m}{NP}-\frac{m}{SP}$

$$=\frac{m}{QP}-\frac{m}{RP}$$

$$=\frac{m}{OP-OQ}-\frac{m}{OP+OR}.$$

If, as before, $OP=d$, and $NS=2l$ we have,

$$\text{Potential at } P=\frac{m(OQ+OR)}{OP^2-OQ^2}=\frac{m\,.\,2l\,.\cos\theta}{OP^2-OQ^2},$$

which for a very short magnet gives—

$$\text{Potential at } P=\frac{M\cos\theta}{d^2}.$$

The same result would have been obtained if the moment of the magnet had been resolved into two components, one along OP, whose value is $M\cos\theta$, and the other perpendicular to OP, whose value is $M\sin\theta$; the potential at P due to the former is $\frac{M\cos\theta}{d^2}$ and that due to the latter is zero, the sum of the two being $\frac{M\cos\theta}{d^2}$. The agreement of this with the previous result justifies us in resolving the magnetic moment into two components. Indeed justification is hardly necessary since magnetic moment is a vector quantity, having direction as well as magnitude, and may therefore be resolved or compounded like all other vector quantities, such as force, velocity, etc.

Equipotential Lines and Surfaces.—A line or surface passing through points having the same potential is an equipotential line or surface. No work is done in carrying a pole along an equipotential line or surface, for the variation of potential along it is zero. Hence by definition of potential, it follows that there is no component of magnetic field along the line or surface, and no force tending to move a magnetic pole along it. From this reasoning it follows that lines of force and equipotential lines always cut each other at right angles, since if they did not there would be a component of the magnetic field acting along the equipotential surface.

Force between Magnets.—The resultant force experienced by a magnet in a uniform field is zero, since the forces on the N and S poles respectively are equal and opposite, the quantities of N and S pole on any magnet being equal. In fact, this absence of resultant force is a most satisfactory proof of the equality of the

two kinds of pole on a magnet. If a bar magnet be floated on a cork in the middle of a large vessel of water, it will experience a couple rotating it into the magnetic meridian, but the magnet will not move from the middle of the vessel, showing that there is no resultant force acting on it.

In the neighbourhood of another magnet the field is not uniform, and in general there will be a resultant force.

Consider the two short magnets NS and N′S′ in Fig. 13. The field at S′ due to NS is $\frac{2M}{x^3}$ where distances are measured from the middle of NS. Hence the force on S′ is $\frac{2M}{x^3} \cdot m'$, where m' is the strength of pole of N′S′.

FIG. 13.

The rate of change of the field due to NS, as we increase x, is

$$\frac{d}{dx}\left(\frac{2M}{x^3}\right) = -\frac{6M}{x^4}.$$

N′S′ being a small magnet of length l, the decrease in field in passing from S′ to $N' = -\frac{6M}{x^4}l$.

$$\therefore \text{ Field at } N' \qquad = \frac{2M}{x^3} - \frac{6M}{x^4}l$$

$$\text{and} \qquad \text{force on } N' \qquad = \left(\frac{2M}{x^3} - \frac{6M}{x^4}l\right)m'.$$

Hence, resultant force on N′S′ being the difference between the forces on S′ and N′,

$$\text{Force on } N'S' = \frac{2M}{x^3} \cdot m' - \left(\frac{2M}{x^3} - \frac{6M}{x^4}l\right)m'$$

$$= \frac{6Mm'l}{x^4}$$

$$= \frac{6MM'}{x^4}.$$

FIG. 14.

In an exactly similar way, we may find the resultant force on N′S′ in the position shown in Fig. 14, but in this case we should note that the field is always parallel to NS, although the variation in field is at right angles to this direction.

Taking y for the distance between the magnets,

$$\text{Force on } N' = \frac{M}{y^3} \cdot m'.$$

$$\left.\begin{array}{r}\text{Rate of variation of field}\\ \text{in the direction of } y\end{array}\right\}=\frac{d}{dy}\left(\frac{\mathrm{M}}{y^3}\right)=-\frac{3\mathrm{M}}{y^4}.$$

$\therefore$ Decrease in field in passing from N′ to S′ $=-\frac{3\mathrm{M}}{y^4}l$,

$$\text{and field at S}'=\frac{\mathrm{M}}{y^3}-\frac{3\mathrm{M}}{y^4}l.$$

$$\text{Force on S}'=\left(\frac{\mathrm{M}}{y^3}-\frac{3\mathrm{M}}{y^4}l\right)m'.$$

$$\text{Resultant force on N}'\text{S}'=\frac{\mathrm{M}}{y^3}m'-\left(\frac{\mathrm{M}}{y^3}-\frac{3\mathrm{M}}{y^4}l\right)m'$$

$$=\frac{3\mathrm{MM}'}{y^4},$$

and is in a direction parallel to NS.

This force on the magnet N′S′ must not be confused with the couple acting on it, the value of which is $\frac{\mathrm{MM}'}{y^3}$, which tends to rotate it into parallelism with NS, but not to give it a motion of translation. The existence of the resultant force explains the following apparent paradox: If the two magnets NS and N′S′ be placed on a floating platform, there is a couple acting on NS due to the presence of N′S′, the value of which is $\frac{2\mathrm{M}'\mathrm{M}}{y^3}$, and N′S′ at the same time experiences a couple $\frac{\mathrm{MM}'}{y^3}$ due to the presence of NS. Since these couples are not equal and opposite, it would at first sight appear that there is a resultant couple acting on the platform due to the interaction of the magnets, which would make the platform rotate continuously. Such a rotation would involve the continuous expenditure of energy without any corresponding supply, which is contrary to experience. But the fallacy consists in neglecting the force of translation $\frac{3\mathrm{MM}'}{y^4}$ acting upon the magnet N′S′ at right angles to the line joining the magnets, which is equivalent to a couple $\frac{3\mathrm{MM}'}{y^4}\times y$ $=\frac{3\mathrm{MM}'}{y^3}$. An inspection of Fig. 14 shows us that this couple would produce an anti-clockwise rotation, while the former would produce a clockwise rotation, so that the difference, or $\frac{2\mathrm{MM}'}{y^3}$, is the resultant couple in an anti-clockwise direction. This is equal and opposite to the couple on NS, which is clock-

wise, so that the two are in equilibrium and the paradox disappears.

Field due to Small Magnet.—The field at the point P due to a very short magnet may be found by resolving the moment M along OP, and at right angles to OP. The former component will produce a field $\frac{2M\cos\theta}{r^3}$ represented by the vector PQ (Fig. 15), and the latter component, the field $\frac{M\sin\theta}{r^3}$ represented by PT. The resultant field is

$$PR=\sqrt{PQ^2+PT^2}=\frac{M}{r^3}\sqrt{4\cos^2\theta+\sin^2\theta}=\frac{M}{r^3}\sqrt{1+3\cos^2\theta},$$

and its inclination to the line OP is RPQ. Now,

$$\tan RPQ=\frac{PT}{PQ}=\frac{\sin\theta}{2\cos\theta}=\tfrac{1}{2}\tan\theta.$$

Hence to find the direction of the resultant field at any point P, make the angle QPA$=\theta=$POV and drop perpendicular AQ

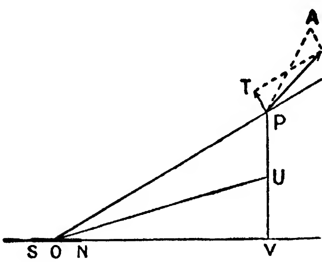


FIG. 15.

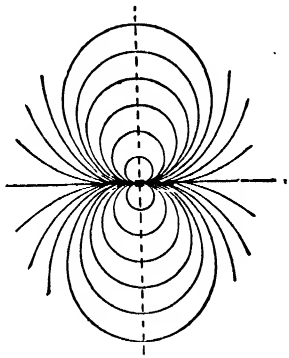


FIG. 16.

upon PQ (Fig. 15). Bisect AQ in R and join PR. PR is then the direction of the field at P. Or if the perpendicular PV be drawn and bisected at U, angle UOV is equal to the angle RPQ made between the direction PR of the field and that of the radius vector OP, and it is the same at all points along OP. By drawing the direction by means of short lines at a number of points along OP and again for a number of different radii, the direction of the field at a number of points is known, and the lines of force may be drawn with fair accuracy.

In Fig. 16, the lines of force for an extremely small magnet have been drawn in this way.

Magnetic Elements.—Great importance attaches to an accurate

knowledge of the condition of the magnetic field due to the earth, both to the navigator for practical purposes, and to the investigator who attempts to describe and account for the magnetic state of the earth. This implies a knowledge, at every instant, of the magnitude and direction of the field at every place, but it is much more convenient to represent the field at any place by means of certain elements or components, than to express it in terms of the resultant field and its direction. For the purpose of representation we choose those elements that lend themselves most readily to experimental determination. These are :—the *Declination* or angle which the magnetic meridian makes with the geographical meridian, α (Fig. 17) ; the *Horizontal Component* of the earth's magnetic field, H, and the *Dip* or angle θ, which the resultant field makes with the horizontal, and it will be seen that when these elements are known the field is completely determined. But it must be borne in mind that the field is always changing, so that the elements undergo variations ; these will be considered later.

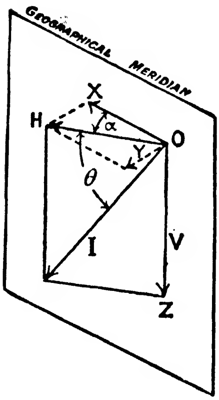


FIG. 17.

Magnetic Meridian.—We have already seen that a suspended magnetic needle sets in a certain direction, approximately N and S. A vertical plane passing through the magnetic axis of a freely suspended needle is called the magnetic meridian. This does not in general coincide with the geographical meridian, and its position may be roughly determined by observing the direction in which a compass needle will set ; but since the magnetic axis of the needle may not coincide with its axis of symmetry or geometric axis, the needle should always be turned over after the first observation has been made, and suspended from the other side. If the geometric axis makes an angle with the magnetic axis, the needle will point E or W of magnetic north, but on suspending it with its other face upwards it will make the same angle on the other side of magnetic north. The direction of the magnetic axis and of the magnetic meridian will then be found by bisecting the angle between the two positions of the axis of symmetry. In Fig. 18, NS is the magnetic meridian. For many purposes the prismatic compass (Fig. 19) may be usefully employed to find approximately the direction of the magnetic meridian. Some distant object in a known geographical direction is sighted by means of the slot S and the wire S′. By

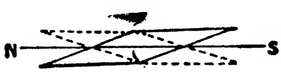


FIG. 18.

means of the right-angled prism shown in section at P the position of the image of the wire S′ may be read upon the scale of the compass card. The magnetic direction of the distant object being thus found, and its geographical direction being known, the position of the magnetic meridian is determined.

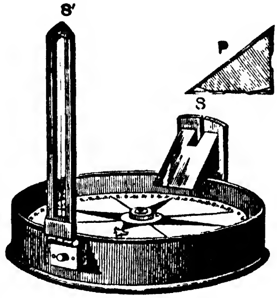


FIG. 19.

Declination.—The angle between the magnetic meridian and the geographical meridian is usually known as the magnetic declination; for nautical purposes it is called the Variation of the Compass, meaning its variation from a true north-and-south direction. In Fig. 17 the plane containing H, I and V is the magnetic meridian, and consequently the angle α is the declination. In the last described experiment the declination is found, but for its more accurate determination the Kew magnetometer, which will be described later, is employed.

Dip.—A perfectly freely suspended magnet would not in general set in a horizontal direction, but along the line of the greatest strength of field. This is represented by I in Fig. 17, and the angle θ between it and the horizontal is called the *magnetic dip.* A perfectly freely suspended magnet is of course an ideal which is unattainable, since the mechanical support must influence the angle at which the needle will set; in fact, a compass needle is deliberately, although perhaps unconsciously, placed in its suspension in such a way that it sets horizontally, and any tendency to dip is neutralised by suspending it from a point which is not its centre of gravity, so that the result of all the forces acting on it is to cause it to remain horizontal. If, however, a needle be mounted on a fine straight axle resting on horizontal knife-edges, and if the axle pass through the centre of gravity of the needle, it can rotate in a vertical plane, and if this plane coincide with the magnetic meridian the needle will set with its magnetic axis along I, Fig. 17, and its inclination to the horizontal will be the dip. The experimental determination of the dip will be described later.

So far we have only determined the direction in space of the resultant field I; if, then, we can find its magnitude or that of any component of it in a known direction, the field becomes completely determined at that particular locality. By far the most convenient component to determine is H, the horizontal component, and then knowing this, the vertical component V may be calculated. For from Fig. 17,

$$\frac{V}{H}=\tan\theta$$

and again, the total intensity I may be found,

$$I^2=H^2+V^2.$$

For some purposes it is convenient to refer the earth's field to three rectangular axes: OX, a horizontal line in the geographical meridian, true N and S; OY, a horizontal line perpendicular to the geographical meridian, true E and W; and OZ a vertical line.

Then taking X, Y and Z as the components of the earth's magnetic field in these directions,

$$\begin{aligned} X&=H\cos\alpha=I\cos\theta\cos\alpha\\ Y&=H\sin\alpha=I\cos\theta\sin\alpha\\ Z&=V=I\sin\theta. \end{aligned}$$

Vibrating Magnet.—Although there are several methods of determining H, that which is most frequently employed is the magnetometric method, its chief advantage over other methods being that it does not involve the use or measurement of electric currents. We have seen on page 7 how the ratio of M to H may be determined for a given bar magnet, in terms of the deflection of a suspended magnetic needle, at a given distance from the magnet. The absolute values of M and H, however, cannot be determined by the magnetometer, but only their ratio. A further experiment is required, to give us some other relation between M and H.

Whenever a suspended magnet makes an angle θ with its position of equilibrium, a couple MH sin θ acts on it (see p. 5) which tends to restore it to that position. Thus the magnet must have an angular acceleration, and we may express the couple acting on it as the product of its moment of inertia I and the angular acceleration $\frac{d^2\theta}{dt^2}$. Then $I\frac{d^2\theta}{dt^2}+MH\sin\theta=0$, since the algebraic sum of the couples acting on it must be zero, the effect of the forces due to the friction, etc., in this case being negligible. Further if the value of θ is never more than a few degrees, the angle itself in circular measure may be taken instead of its sine, and our equation is therefore—

$$I\frac{d^2\theta}{dt^2}+MH\theta=0.$$

As this type of equation will frequently occur, we will proceed to solve it. First write it in the form—

$$\frac{d^2\theta}{dt^2}+k^2\theta=0,$$

where k^2 is substituted for $\frac{MH}{I}$.

An equation of this type can be solved in the form $\theta=\epsilon^{at}$. Differentiating this last equation twice, we have—

$$\frac{d^2\theta}{dt^2}=a^2\epsilon^{at},$$

and substituting the values for θ and $\frac{d^2\theta}{dt^2}$ in the equation, we have—

$$a^2\epsilon^{at}+k^2\epsilon^{at}=0$$

$$\therefore\ a^2=-k^2,\text{ and, }a=\pm\sqrt{-k^2}$$
$$=\pm k\sqrt{-1}.$$

Thus there are two particular solutions—

$$\theta=A\epsilon^{k\sqrt{-1}t},\ \text{ and, }\ \theta=B\epsilon^{-k\sqrt{-1}t},$$

and the most general equation to the motion of the needle is—

$$\theta=A\epsilon^{k\sqrt{-1}t}+B\epsilon^{-k\sqrt{-1}t}$$

where A and B are two constants that can be determined from the conditions of the problem. Thus, if the time be reckoned from the instant at which the needle is in the direction of the meridian, $\theta=0$ when $t=0$,

$$\therefore\ A+B=0,\text{ or, }A=-B,$$

and the equation may be written—

$$\theta=A(\epsilon^{k\sqrt{-1}t}-\epsilon^{-k\sqrt{-1}t}).$$

Again, the angular velocity of the needle is—

$$\frac{d\theta}{dt}=k\sqrt{-1}A(\epsilon^{k\sqrt{-1}t}+\epsilon^{-k\sqrt{-1}t}),$$

and if this is ω when t and θ are zero—

$$A=\frac{\omega}{2k\sqrt{-1}},$$

$$\therefore\ \theta=\frac{\omega}{k}\left(\frac{\epsilon^{k\sqrt{-1}t}-\epsilon^{-k\sqrt{-1}t}}{2\sqrt{-1}}\right).$$

The term in brackets is the well-known exponential form of the sine of the angle kt, and so writing $\frac{\omega}{k}=\theta_0$, we have—

$$\theta=\theta_0\sin kt.$$

Thus we see that the needle executes simple harmonic oscillations,

one complete oscillation occurring in time $\frac{2\pi}{k}$; for, on increasing t by this value we get—

$$\theta=\theta_0 \sin k\left(t+\frac{2\pi}{k}\right)=\theta_0 \sin (kt+2\pi)=\theta_0 \sin kt$$

and the motion is repeated after intervals of time $\frac{2\pi}{k}$.

Calling the periodic time T, we have $T=\frac{2\pi}{k}$, and remembering that $k^2=\frac{MH}{I}$, we see that $T=2\pi\sqrt{\frac{I}{MH}}$.

Determination of M and H.—The moment of inertia I of the magnet may be found from its mass and linear dimensions. In the case of a rectangular magnet, $I=\text{mass}\times\frac{\text{length}^2+\text{breadth}^2}{12}$, and for a cylindrical magnet, $I=\text{mass}\times\left(\frac{\text{length}^2}{12}+\frac{\text{radius}^2}{4}\right)$. The time of oscillation is observed by suspending the magnet in the locality previously occupied by the needle in the magnetometer experiment and observing the time of a number of swings. Then from our equation we have $MH=\frac{4\pi^2 I}{T^2}$, and the magnetometer experiment gave us $\frac{M}{H}$. Combining these we have, $MH\times\frac{M}{H}=M^2$, or $MH\div\frac{M}{H}=H^2$, so that both M and H are now determined in absolute measure.

Comparison of Fields by Vibration.—It may be noticed that $H=\frac{4\pi^2 I}{MT^2}$, and hence, that if the same magnet be employed on different occasions, I and M will be the same, so that $\frac{H_1}{H_2}=\frac{T_2^2}{T_1^2}$, or if n be the number of oscillations made by the magnet in a given time, $\frac{n_1}{n_2}=\frac{T_2}{T_1}$, so that $\frac{H_1}{H_2}=\frac{n_1^2}{n_2^2}$. This gives rise to a method of comparing the strengths of magnetic field at two given times or places; for if the same magnet be allowed to oscillate on the two occasions and the number of oscillations made in equal times observed, the ratio of the two field strengths is known. But it must be noticed that the fibre used to suspend the magnet must be as nearly as possible torsionless, and further, the magnet must be carefully preserved from ill treatment, mechanical or

thermal, or its magnet moment will not be the same on the two occasions.

To determine the time of swing, the magnet is suspended in a vibration box (Fig. 20) by means of a silk support, the magnet being held in a double loop as shown. It is advisable to suspend some bar of approximately the same weight as the needle before placing the magnet in position, in order that any torsion in the thread may be removed. If the suspension head be turned so that the solid at rest lies approximately in the meridian, then on replacing it by the magnet the suspending fibre will be very nearly free from torsion. The position of equilibrium of the magnet may be marked upon the front and back glass walls of the box and the magnet then given a small oscillation. At the instant of passing the equilibrium position in one direction, the time by the chronometer is noted or the stop-watch is started. At the passage across the equilibrium position in the same direction after fifty or one hundred swings the time is again noted, and the time for one oscillation may then be found by division.

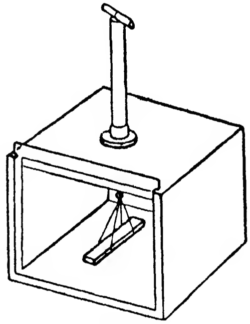


Fig. 20.

Equivalent Length of Magnet.—Employing the magnetometer p. 10, the deflection may be found as there described, and the mean value of $\frac{M}{H}$ found from the expression $\frac{M}{H}=\frac{d^3}{2}\tan\theta$ for the "A" position of Gauss, or $\frac{M}{H}=d^3\tan\theta$ for the "B" position. The use of these approximate formulæ is justified if l^2 is negligible in comparison with d^2, or $\frac{l^2}{d^2}$ is a less percentage of unity than the percentage error introduced in making the observations of deflection. The difficulty arises that if d is made very great, the deflection may be so small that the error in its measurement is considerable. Hence it is desirable to employ the more exact formulæ—

$$\frac{M}{H}=\frac{(d^2-l^2)^2}{2d}\tan\theta,\ \text{and,}\ \frac{M}{H}=(d^2+l^2)^{\frac{3}{2}}\tan\theta,$$

but unfortunately l is unknown, the poles being distributed over a large surface of the magnet. The effective value of L or $2l$, the length of the magnet, may be found from two measurements of $\frac{d^3}{2}\tan\theta$, the approximate value of $\frac{M}{H}$, for two distances from

the suspended needle, and may then be applied as a correction to obtain the true value of $\frac{M}{H}$.

Calling $\frac{M}{H}$ the true value, $\left(\frac{M}{H}\right)_1$ the value found for distance d_1, by using the approximate formula, and $\left(\frac{M}{H}\right)_2$ that for distance d_2, we see that—

$$\frac{M}{H}=\frac{(d_1^2-l^2)^2}{2d_1}\tan\theta_1=\frac{d_1^4\left(1-\frac{l^2}{d_1^2}\right)^2}{2d_1}\tan\theta_1$$

$$=\left(1-\frac{l^2}{d_1^2}\right)^2\frac{d_1^3}{2}\tan\theta_1$$

$$=\left(1-\frac{l^2}{d_1^2}\right)^2\left(\frac{M}{H}\right)_1$$

$$=\left(\frac{M}{H}\right)_1\left(1-\frac{2l^2}{d_1^2}+\frac{l^4}{d_1^4}\right)$$

Since $\frac{l^2}{d_1^2}$ is small, we may reasonably say that $\frac{l^4}{d_1^4}$ is negligible, and remembering that $L=2l$,

$$\frac{M}{H}=\left(\frac{M}{H}\right)_1\left(1-\frac{L^2}{2d_1^2}\right).$$

Similarly, $$\frac{M}{H}=\left(\frac{M}{H}\right)_2\left(1-\frac{L^2}{2d_2^2}\right),$$

$$\therefore\ \left(\frac{M}{H}\right)_1-\left(\frac{M}{H}\right)_1\frac{L^2}{2d_1^2}=\left(\frac{M}{H}\right)_2-\left(\frac{M}{H}\right)_2\frac{L_2}{2d_2^2}$$

and, $$\frac{L^2}{2}=\frac{\left(\frac{M}{H}\right)_1-\left(\frac{M}{H}\right)_2}{\left(\frac{M}{H}\right)_1\frac{1}{d_1^2}-\left(\frac{M}{H}\right)_2\frac{1}{d_2^2}}.$$

The quantity on the right being determined from measurements at two distances, we know the correction $\left(1-\frac{L^2}{2d^2}\right)$ to be applied to the approximate formula to obtain the true value—

thus, $$\frac{M}{H}=\left(\frac{M}{H}\right)_1\left(1-\frac{L^2}{2d_1^2}\right)=\left(\frac{M}{H}\right)_2\left(1-\frac{L^2}{2d_2^2}\right).$$

The quantity $\frac{L^2}{2}$ may be taken as one constant P, and if in addition $A_1=\left(\frac{M}{H}\right)_1$ and $A_2=\left(\frac{M}{H}\right)_2$

$$P=\frac{A_1-A_2}{\frac{A_1}{d_1^2}-\frac{A_2}{d_2^2}}, \quad \text{and,} \quad \frac{M}{H}=\left(\frac{M}{H}\right)_1\left(1-\frac{P}{d_1^2}\right).$$

For the " B " position of Gauss the correction may be applied in a similar manner—

$$\frac{M}{H}=(d^2+l^2)^{\frac{3}{2}} \tan \theta=d^3\left(1+\frac{l^2}{d^2}\right)^{\frac{3}{2}} \tan \theta$$

$$=\left(\frac{M}{H}\right)_1\left(1+\frac{l^2}{d_1^2}\right)^{\frac{3}{2}}$$

$$=\left(\frac{M}{H}\right)_2\left(1+\frac{l^2}{d_2^2}\right)^{\frac{3}{2}}.$$

$$\therefore \left(\frac{M}{H}\right)_1\left(1+\frac{3}{2}\cdot\frac{l^2}{d_1^2}+\frac{\frac{3}{2}\cdot\frac{1}{2}}{1\,.\,2}\cdot\frac{l^4}{d_1^4}+\ \ldots\right)$$

$$=\left(\frac{M}{H}\right)_2\left(1+\frac{3}{2}\cdot\frac{l^2}{d_2^2}+\frac{\frac{3}{2}\cdot\frac{1}{2}}{1\,.\,2}\cdot\frac{l^4}{d_2^4}+\ \ldots\right).$$

Neglecting $\frac{l^4}{d^4}$ and higher powers of $\frac{l}{d}$, we have as before—

$$\left(\frac{M}{H}\right)_1\left(1+\frac{3}{2}\cdot\frac{l^2}{d_1^2}\right)=\left(\frac{M}{H}\right)_2\left(1+\frac{3}{2}\cdot\frac{l^2}{d_2^2}\right)$$

Substituting L for $2l$ we have—

$$\left(\frac{M}{H}\right)_1\left(1+\frac{3}{8}\cdot\frac{L^2}{d_1^2}\right)=\left(\frac{M}{H}\right)_2\left(1+\frac{3}{8}\cdot\frac{L^2}{d_2^2}\right)$$

$$\therefore \tfrac{3}{8}L^2=\frac{\left(\frac{M}{H}\right)_2-\left(\frac{M}{H}\right)_1}{\left(\frac{M}{H}\right)_1\cdot\frac{1}{d_1^2}-\left(\frac{M}{H}\right)_2\frac{1}{d_2^2}}=P.$$

$$\therefore \frac{M}{H}=\left(\frac{M}{H}\right)_1\left(1+\frac{P}{d_1^2}\right)=\left(\frac{M}{H}\right)_2\left(1+\frac{P}{d_2^2}\right).$$

Thus, from the deflection produced at any two distances of the magnet, the true value of $\frac{M}{H}$ may be found ; or, if it is desired, the equivalent length L of the magnet can be obtained. If this

be found for any magnet, future determinations of $\frac{M}{H}$ may be made by observing the deflection for one distance of the deflecting magnet.

The Kew Magnetometer.—The form of needle used in the Kew

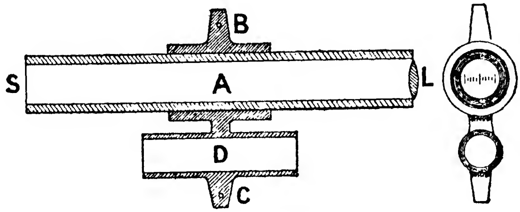


Fig. 21.
(*From Watson's "Textbook of Physics."*)

pattern of magnetometer is shown in Fig. 21. It consists of a steel tube A having a fine transparent scale S at one end and a lens L at the other, the scale being at the principal focus of the lens. The magnet is thus a collimator, and when the telescope is focused for infinity and placed co-axially with the magnet, an image of the fine scale will be seen in the focal plane of the telescope. If the suspension fibre be freed from torsion and the body of the magnetometer (Fig. 22) rotated until the image of the middle division of the scale coincides with the cross wire of the telescope, the azimuth of the telescope, as indicated by the horizontal circular scale, gives the direction of the geometric axis of the magnet. The magnet is then turned over and the position of the geometric axis with reference to the horizontal scale again determined. The mean of the two positions is the azimuth

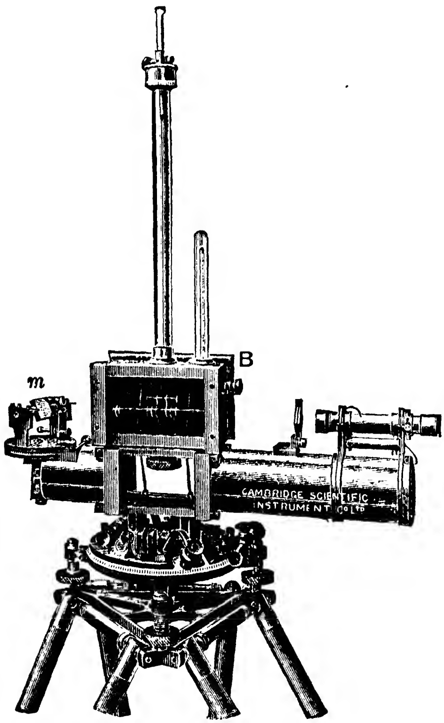


Fig. 22.

of the magnetic meridian upon this scale. If the azimuth of the geographical meridian also be found, the difference between the two gives us the magnetic declination. The geographical meridian is found by observing the image of the sun produced by the mirror m in passing the cross wire of the telescope, it having been previously adjusted so that its axis is horizontal, and the plane in which the normal travels as the mirror is turned contains the optic axis of the telescope. From the observed time of the sun's passing the cross wire, knowing the longitude of the place of observation and the equation of time, the direction of the sun at the time of observation is known, and thus the direction of true N and S also.

The period of oscillation of the magnet may now be found with the arrangement just described. The magnet is given a small oscillation, and the time for 100 transits of the image of the middle scale division across the cross wire of the telescope in one direction, either from left to right, or right to left, is observed. This gives the time for 100 oscillations, from which the time of one oscillation must be found. This must be corrected, for the fact that the fibre, although very fine, yet exerts some controlling couple on the magnet, and therefore shortens the period of oscillation. If the suspension head be rotated through 90°, an angular deviation of a radian is produced and may be observed; then $\frac{\pi}{2}-a$ is the twist in the suspension, and

$$c\left(\frac{\pi}{2}-a\right)=\text{MH}\sin a=\text{MH}a,$$

where c is the couple exerted by the fibre for one radian twist, and a is the very small deflection produced by 90° rotation of the torsion head.

When the magnet is at an angle θ to the magnetic meridian, the restoring couple is now $(\text{MH}+c)\theta$ instead of $\text{MH}\theta$, and the time of oscillation found on page 24, will therefore be $2\pi\sqrt{\frac{\text{I}}{\text{MH}+c}}$ instead of $2\pi\sqrt{\frac{\text{I}}{\text{MH}}}$. It follows that from the observed time of swing we have really obtained

$$\frac{4\pi^2\text{I}}{\text{T}^2}=\text{MH}+c=\text{MH}+\frac{\text{MH}a}{\frac{\pi}{2}-a}=\text{MH}\left(1+\frac{a}{\frac{\pi}{2}-a}\right)$$

$$\therefore\ \text{MH}=\frac{4\pi^2\text{I}}{\text{T}^2\left(1+\frac{a}{\frac{\pi}{2}-a}\right)},$$

and we see that the square of the observed time of swing must be multiplied by the factor $\left(1+\frac{a}{\frac{\pi}{2}-a}\right)$ in order to obtain the corrected value of MH.

Two further corrections must be applied: one for the fact that the magnetic moment of the magnet changes with temperature, and the other for the fact that, being in the earth's magnetic field, its magnetic moment is greater than when, as in the deflection experiment, it is in an E and W direction. The first of these corrections is made by reducing the moment to that at 0° C. by means of the factor $\{1+q(t-t_0)\}$; thus $M_0=M_t\{1+q(t-t_0)\}$, the magnetic moment decreasing with rise of temperature: q must be found by a previous experiment for each individual magnet. The second correction is applied by means of a factor in which it is assumed that the alteration in magnetic moment is proportional to the field in which the magnet is situated. This assumption is justified if the field is small, and further, the change in magnetic moment is proportional to the volume of the magnet, and depends on the position of the magnet in relation to the field, and the material of which the magnet is made (see Chapter IX); thus if M_0 is the moment in zero field or whenever the magnet is situated at right angles to the field, and M that when parallel to the field, $M=M_0+\alpha VH$, where α is some constant depending on the nature of the material of the magnet. V is also constant, so calling $\alpha V=b$ we have $M=M_0+bH$.

$$\therefore\ MH=M_0H+bH^2=M_0H\left(1+b\frac{H}{M_0}\right).$$

$\frac{H}{M_0}$ is always very small, and hence, when the whole quantity $b\frac{H}{M_0}$ is small,

$$M_0H=MH\left(1-b\frac{H}{M_0}\right).$$

$\frac{H}{M_0}$ is known from the deflection experiment, and b is a constant for a magnet of any given size and material, so that the quantity MH as found from the vibration experiment may be corrected by means of the factor $\left(1-b\frac{H}{M_0}\right)$.

The square of the time of oscillation may thus be corrected for torsion of fibre, temperature and alteration of moment due to the magnetic field, by a single factor, and

$$T_0^2=T^2\left\{1+\frac{a}{\frac{\pi}{2}-a}-q(t-t_0)+b\frac{H}{M_0}\right\}.$$

The moment of inertia I of the magnet and carrier may be found by adding a body of known moment of inertia and re-determining the time of oscillation. A brass cylinder which just fits into the space D of the carrier (Fig. 21) is employed. If I_1 is the moment of inertia of the cylinder, $I_1=m\left(\frac{l^2}{12}+\frac{r^2}{4}\right)$, and the time of vibration is now $T_1=2\pi\sqrt{\frac{I+I_1}{MH}}$.

$$T^2=4\pi^2\cdot\frac{I}{MH},\ \text{and},\ T_1^2=4\pi^2\frac{I+I_1}{MH},$$

$$\therefore \frac{I+I_1}{I}=\frac{T_1^2}{T^2}$$

from which,

$$I=I_1\frac{T^2}{T_1^2-T^2}.$$

I being found in this way, the value may be used for any number of vibration experiments with the same magnet, since it is a mechanical constant and is independent of the magnetic condition of the magnet.

To perform the deflection experiment for the determination of $\frac{M}{H}$, the box B (Fig. 22) is removed, and a small magnet with mirror attached is suspended by a long fibre. The telescope is focused upon the image of the scale S, Fig. 23, in this mirror, and the collimator magnet of the vibration experiment is placed in the V rest at M upon the carrier, which may be set at different distances from the mirror needle by moving it along the graduated bar XY. In this experiment the deflection of the needle is not observed, as we should then have to apply a correction for the torsion introduced into the fibre when the needle rotates, but instead, the body of the instrument is rotated until the middle division of the scale coincides with the cross wire of the telescope, and this rotation is measured by taking the difference between the various readings upon the horizontal circular scale with and without the presence of the collimator magnet M. In this case, field due to magnet NS is $\frac{2M}{d^3}$, and couple on needle is $\frac{2M}{d^3}m$, where m is the magnetic moment of the needle.

Restoring couple due to earth's field H, is $Hm \sin \theta$.

$$\therefore \frac{2M}{d^3}m = Hm \sin \theta, \text{ or } \frac{M}{H} = \frac{d^3}{2} \sin \theta.$$

Readings for correcting as on pp. 27 and 31 are then made.

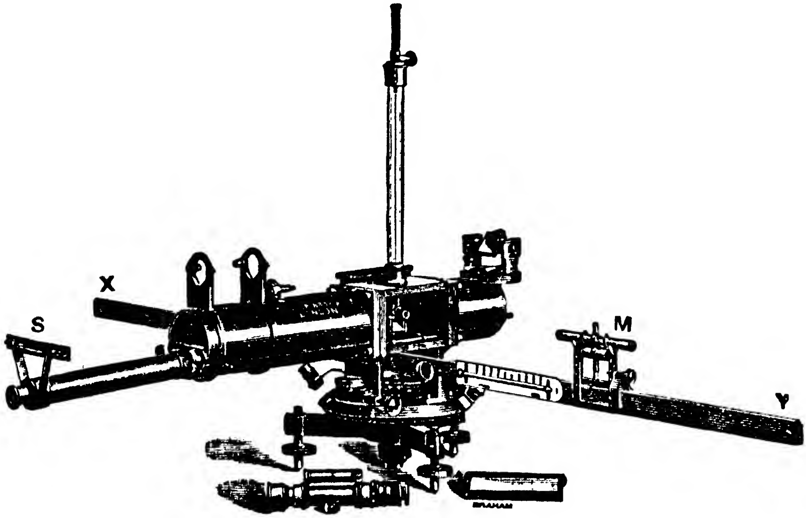


FIG. 23.

Schuster's Method of determining H.—For the accurate determination of H at an observatory, Sir A. Schuster [1] describes the method of comparing it with the uniform field due to the current in a pair of coils (p. 53). If the coil AB (Fig. 24) carries current i, the magnetic field Fi due to it can be calculated. The coil AB is rotated about a vertical axis until the component $Fi \cos \alpha$ is equal and opposite to H. The component $Fi \sin \alpha$ then sets the suspended magnet NS at right angles to the direction of H,

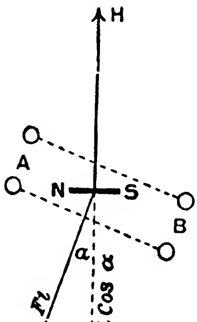


FIG. 24.

then, $$H = Fi \cos \alpha.$$

By choosing i so that Fi has nearly the same value as H, the angle α is made small, and $\cos \alpha$ varies slowly. F is calculated from the dimensions of the coil AB (p. 53), and α is observed, a check being obtained by observing the symmetrical position for AB, for which the magnet NS is reversed in direction. In this way H may be measured to

[1] A. Schuster, *Terrestrial Magnetism*, xix. March, 1914.

a few parts in 100,000, the observations requiring only a few minutes to make.

Determination of Dip.—The Kew pattern of dip circle is shown in Fig. 25. The dip needle itself is a thin steel magnet AB provided with a fine steel axle which rests on agate knife-edges shown at K, K′.

It is carried by the V supports L, L′ which may be raised and lowered by turning the milled head E. In making a reading, the observer must raise the needle and lower it repeatedly, in order to bring the axle constantly to the centre of the circular

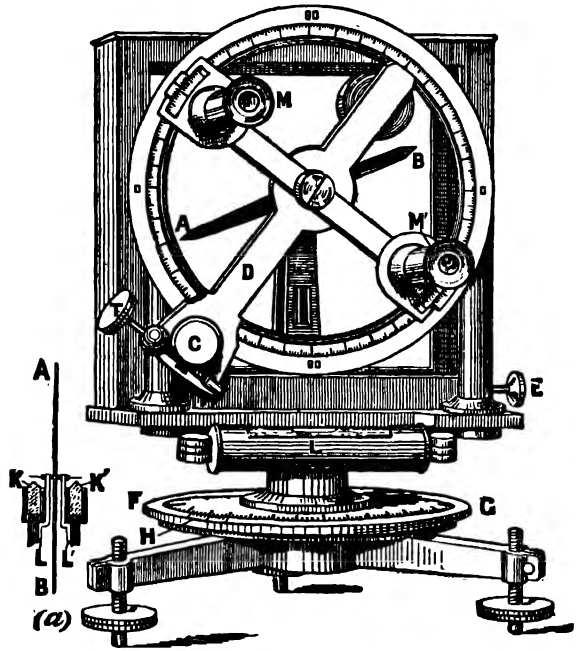


Fig. 25.
(*From Watson's "Textbook of Physics."*)

scale, since as the needle swings it rolls upon the axle and travels from the centre of the scale. The body of the instrument may be rotated about a vertical axis and its azimuth read upon the horizontal scale H. The position of the needle with reference to the vertical or actual dip circle is found by rotating the arm which carries the microscopes M, M′ until the cross wires appear to coincide with the tips of the needle, the verniers being then read.

To begin observations, the instrument is levelled and then rotated about its vertical axis until the needle sets vertically,

i.e. reads 90°—90°. The plane in which the needle rotates is then at right angles to the magnetic meridian. For, let the plane AB (Fig. 26) be the plane of rotation of the needle, and let it make angle δ with the magnetic meridian HB.

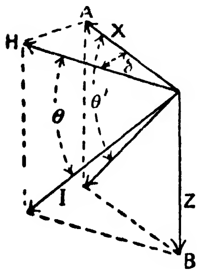


FIG. 26.

Resolving the magnetic field I into three components X and Z in this plane, and Y at right angles to it, X being horizontal and Z vertical,

$$X = H \cos\delta = I \cos\theta \cos\delta$$
$$Y = H \sin\delta = I \cos\theta \sin\delta$$
$$Z = I \sin\theta.$$

Y, being in the direction of the axle, will not exert any turning moment about it, and the needle will therefore set along the resultant of X and Z. If we call θ' the angle that this resultant makes with the horizontal,

$$\tan\theta' = \frac{Z}{X} = \frac{I \sin\theta}{I \cos\theta \cos\delta} = \frac{\tan\theta}{\cos\delta}$$

θ is the true dip given by $\tan\theta = \dfrac{I \sin\theta}{I \cos\theta} = \dfrac{Z}{H}$, and is for the present treated as a constant quantity at the given locality.

$$\text{If now, } \theta' = 90°, \ \tan\theta' = \infty$$
$$\therefore \cos\delta = 0, \quad \text{and } \delta = 90°$$

Thus the plane of rotation of the dipping needle is at right angles to the meridian, and on rotating the instrument about its vertical axis through 90° as determined by the horizontal circle, the plane of rotation of the needle will be brought into the meridian, and the measured dip will then be the true dip. Note that if $\delta = 0$, $\cos\delta = 1$, and $\tan\theta' = \tan\theta$.

Another method of using the dip circle consists in measuring the dip in any two positions of the circle the angle between which positions is 90°. The instrument is clamped to its vertical axis and the position upon the scale FG noted, and the dip in this position is found.

If δ is the angle between the plane of rotation of the needle and the magnetic meridian, and θ_1 the observed dip, we have already seen that $\tan\theta_1 = \dfrac{\tan\theta}{\cos\delta}$. On now rotating the instrument through 90°, and again observing the dip θ_2, we have—

$$\tan\theta_2 = \frac{\tan\theta}{\cos(\delta + 90°)} = \frac{\tan\theta}{-\sin\delta}.$$

$$\therefore \frac{1}{\tan^2\theta_1} = \frac{\cos^2\delta}{\tan^2\theta}, \quad \text{and,} \quad \frac{1}{\tan^2\theta_2} = \frac{\sin^2\delta}{\tan^2\theta}.$$

Adding, we get—

$$\frac{1}{\tan^2\theta}=\frac{1}{\tan^2\theta_1}+\frac{1}{\tan^2\theta_2}, \text{ or, } \operatorname{Cot}^2\theta=\operatorname{Cot}^2\theta_1+\operatorname{Cot}^2\theta_2.$$

Errors in determining Dip.—In measuring the dip there are several errors to be eliminated, and if these are small, the mean of the following readings will give the true dip :—

(i) The positions of the two ends of the needle are read, in order to correct for the fact that the centre of rotation of the needle may not be at the centre of the vertical circle (Fig. 27 (i)).

(ii) The instrument is rotated through 180° about the vertical axis, and the two previous readings repeated, since the 0°—0° line of the vertical scale may not be horizontal and the apparent dip, if too great in the first position, will be too small by an equal amount in the second position. The zero line after rotating the instrument through 180° is 0′—0′ (Fig. 27 (ii)).

(iii) The needle is turned over on its bearings and the previous four readings repeated, because the magnetic axis of the needle

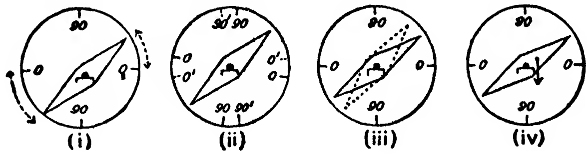


FIG. 27.

may not coincide with its geometric axis as described on p. 20 (Fig. 27 (iii)).

(iv) The magnet is remagnetised in the opposite direction, so that the end which dipped previously now points upwards. This is the only way of correcting for the fact that the axis of rotation may not pass through the centre of gravity of the needle, a small couple due to gravity causing the needle to rotate from the position of true dip (Fig. 27 (iv)). The previous eight readings are repeated and the mean of the whole sixteen taken as the true dip. The individual readings should never differ by more than a degree from the mean, if the instrument is properly constructed.

Magnetic Maps.—The three magnetic elements, Declination, Dip and Horizontal Intensity, having been observed at a great number of stations, the question arises as to how the results may be represented to the greatest advantage. Many methods have been employed, but the most frequent is to draw lines upon a map, passing through all points for which one of the magnetic elements has a common value. Thus three maps are required, one for the representation of each element, or the three may be

represented upon one map. Lines passing through points having the same value of the declination are called *Isogonal lines*, those passing through points for which the dip is the same are *Isoclinal lines*, and *Isodynamic lines* are those passing through points for which the horizontal intensity is the same.

In Fig. 28 the isogonals for the year 1947 are represented on a map of the world drawn on Mercator's projection. They converge towards four points upon the earth's surface, namely, the two geographical poles, and two other points, called the magnetic poles. There are two chief *agonic* lines or lines of no declination, at all points of which the compass points towards the geographical poles. One of these agonic lines passes from the magnetic north pole to the geographic south pole by way of America and the Atlantic Ocean, and the other from the geographic north pole to the magnetic south pole through Eastern Europe, Arabia, the Indian Ocean and Australia. Along some line joining the magnetic and the geographic north poles, the declination is 180° ; the N pole of the compass points towards the magnetic pole and therefore away from the geographic pole. A similar state of affairs exists between the magnetic and geographic south poles. These facts can be much better realised by drawing the isogonals upon a globe, and cannot be adequately represented upon a plane diagram.

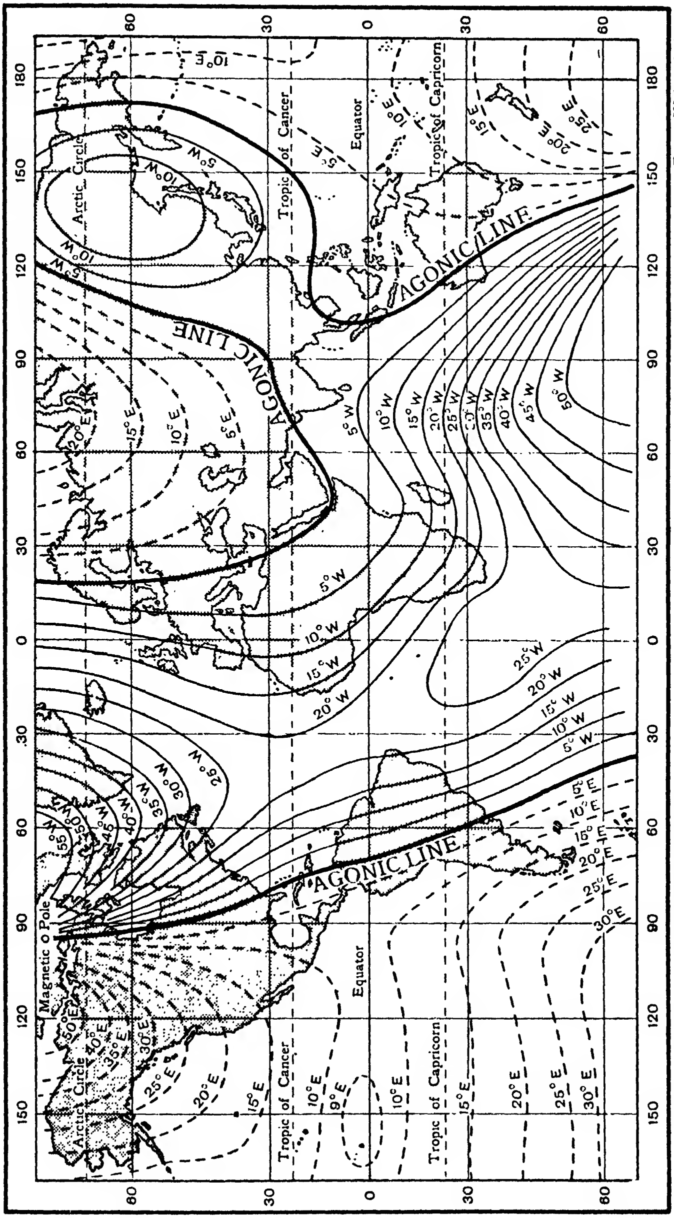
East of the American agonic line the declination is westerly, that is, the compass needle points west of true north ; the isogonals of westerly declination are full lines in Fig. 28. The isogonals of easterly declination are dotted lines and lie west of the American agonic line. It will be seen that the isogonals are far from being regular curves. They reach their greatest irregularity in Eastern Asia, where there is a district surrounded by a looped agonic line within which the declination is westerly. This is called the Siberian Oval.

Instead of isogonals, the lines which indicate the direction of the magnetic meridian are sometimes plotted. These are called lines of magnetic longitude, or Duperrey's lines, and they are more regular than the isogonals. They also differ from the isogonals in converging to only two points—the magnetic poles.

The lines of equal dip, or isoclinals, are much more regular than the isogonals ; they approximate to circles on the sphere, having poies at the magnetic poles. The line of no dip is called the *magnetic equator*. It crosses the geographic equator twice, once in the Atlantic and once in the Pacific Ocean, and lies south of it in the American Hemisphere. The other isoclinals are roughly parallel to the magnetic equator, and therefore correspond to parallels of latitude. The points at which the dip is 90° are the

magnetic poles. The North magnetic pole was reached by Sir James Ross in 1831 and found to be in longitude 96° 43′ W.,

LINES OF EQUAL DECLINATION FOR THE YEAR 1947



Emery Walker Ltd. sc.

Fig. 28.

Based upon British Admiralty Chart No. 2598 with the permission of the Controller of H.M. Stationery Office and of the Hydrographer of the Navy

latitude 73° 31′ N. The South magnetic pole was reached by Sir Ernest Shackleton's expedition on 16th January, 1909, the

magnetic observations being made by Dr. Mawson,[1] who found the dip to be 90° in latitude 72° 25′ S., longitude 155° 16′ E. It will thus be seen that the magnetic poles are not quite at opposite ends of a diameter of the earth. The approximate magnetic axis of the earth makes an angle of about 17° with the axis of rotation.

Theory of Terrestrial Magnetism.—As a first approximation, the earth's field may be represented as that due to a short magnet placed at its centre, whose direction is in the line joining the magnetic poles. To represent consistently the magnetic condition of the earth, we must assume that the pole of this fictitious small magnet which lies under the North magnetic pole, has pole of the kind which we have called S, or South seeking, for it evidently attracts the N pole of a suspended magnet and repels the S pole. For such a magnetic condition, the dip needle would be horizontal at such points as A and B (Fig. 29) and vertical at C and D. The latter correspond to the magnetic poles, and the former lie on the magnetic equator. At a point, P, such that the angle subtended by the arc PC at the centre is α, or the "magnetic latitude" λ is $(90°-\alpha)$, we can easily find the dip. For calling m the magnetic moment of NS, and R the radius of the earth,

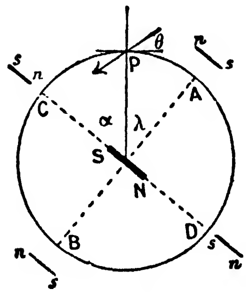


FIG. 29.

$$\text{Component of moment along radius P} = m\cos\alpha$$
$$= m\sin\lambda,$$

and the field at P due to this is $\dfrac{2m\sin\lambda}{R^3}$, and is vertical.

Component of moment perpendicular to radius P is

$$m\sin\alpha = m\cos\lambda,$$

and field at P due to it is $\dfrac{m\cos\lambda}{R^3}$, and is horizontal.

If, then, θ is the dip, $\tan\theta = \dfrac{2m\sin\lambda}{R^3}\cdot\dfrac{R^3}{m\cos\lambda}$

$$=2\tan\lambda.$$

As a rough approximation this is useful, but a glance at Fig. 28 shows that no such simple representation of the earth's magnetic condition is possible.

The fact that the magnetic field of the earth is probably due to magnetisation of the earth's material was first pointed out by

[1] E. Shackleton, "The Heart of the Antarctic"

Dr. Gilbert (1540–1603) of Colchester, who made a model, or terella, of magnetite, and showed that a small suspended needle in the neighbourhood of it, dips as a needle does in the earth's field; but he made the mistake of assuming that the magnetic poles were at the ends of the axis of rotation of the earth, and attributed the declination of the compass to irregularly disposed masses of magnetic material in the earth.

The greatest step forward in the theory of terrestrial magnetism was made by Gauss.[1] By mapping out a closed path upon the earth's surface and resolving the horizontal component of the earth's field along it, he obtained the quantity H cos δ at each point, and by finding the quantity $\int H \cos \delta \,.\, ds$ for short steps ds, taken round the closed path, he found that the result is zero. Hence the magnetic field is not due to an electric current flowing through the curve, that is, there is no vertical current (Chap. VIII). For this purpose he used a triangle, with Göttingen, Milan and Paris, as vertices, and found the above to be true within the error of observation. He also calculated the value of the potential at all points upon the earth in terms of the horizontal field at a limited number of places, and so obtained the values of the total intensity and dip all over the earth. The mathematical discussion is beyond the scope of this book, and the student who is interested in the matter is referred to Gauss's original memoir, or, for a short account, to A. Gray's *Treatise on Magnetism and Electricity*, Vol. I.

Amongst other results, Gauss calculated that the north magnetic pole would be in latitude 73° 35′ N., longitude 95° 39′ W., and the south magnetic pole in latitude 72° 35′ S., longitude 152° 30′ E.; also the magnetic moment of the earth to be about $0{\cdot}33R^3$, where R is its radius.

In Chap. IX we shall see that the magnetic moment of a uniformly magnetised sphere is $\frac{4}{3}\pi R^3 I$, where I is the intensity of magnetisation, and therefore, considering the earth to be a uniformly magnetised sphere, its intensity of magnetisation would appear to be

$$\frac{\cdot 33}{\frac{4}{3}\pi} = \frac{1}{4\pi} \text{ approx.}$$
$$= 0{\cdot}08.$$

The saturation intensity of magnetisation of iron or steel is of the order 1500, and we thus get an idea of the intensity of magnetisation of the earth required to produce its magnetic field. The surface layers of the earth are not capable of so great an intensity of magnetisation as is required by Gauss's theory, so

[1] Gauss, *Allgemeine Theorie des Erdmagnetismus.* Result. d. Magnetischen Vereins, Leipzig. 1839.

that either the interior of the earth is much more highly magnetic than the layers near the surface, or the magnetic field is due to some other cause, such as circular electric currents flowing from east to west. The theory of the earth's permanent magnetic field is far from complete. (See further, p. 46.)

Variation in Magnetic Elements.—The magnetic elements at all points are continually changing, and the change may be resolved into a number of quasi periodic components, together with sudden and irregular changes known as magnetic storms.

(i) *Secular Variation.*—The declination at all points is undergoing a long period change. Records of the declination do not go back far enough for us to compute with accuracy the periodic time of the secular variation, but it is of the order of magnitude of 960 years. In 1580 the declination at London was 11° 15′ E.; in 1600, 5½° E. According to an observation in 1633 it was still 4° 5′ E., and in 1659 it was zero, the compass at London pointing due north. Later observations show a westerly variation, 10½° in 1709, to 24½° in 1820, when it reached its maximum, and has since been diminishing. In the year 1933 it was 11° 52′ W., and it is probable that in 2139 it will again be zero. It was pointed out by Lord Kelvin that the magnetic system is slowly rotating from east to west, making a revolution in 960 years, so that in 960 years the magnetisation lags behind the earth by one rotation. The magnetic north pole describes a small circle of about 17° radius, and the effect of this rotation upon the declination at any fixed point may be seen from Fig. 30.

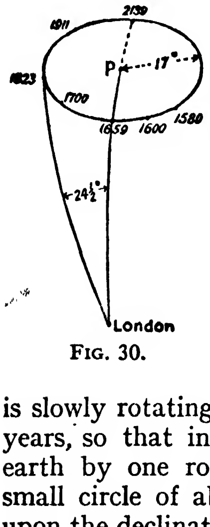


FIG. 30.

(ii) *Annual Variations.*—There is a variation in declination whose periodic time is one year, which occurs simultaneously in opposite directions in the northern and southern hemispheres, the amplitude at London being about 2¼′. The maximum easterly deviation occurs in August, and the westerly in February.

(iii) *Daily Variation.*—Changes in the earth's field having a period of 24 hours are also observed. In Fig. 31 curve A gives the typical variation $\delta\theta$ in the declination in this country. This reaches a limiting position about 4′ east of its mean position just before 8 a.m. and a maximum 5′ west at 1·0 p.m. The variation δH in horizontal intensity (B) reaches a minimum at about 10 a.m. and a maximum at 7 p.m., while the variation $\delta\theta$ in dip (C) reaches its maximum at 11 a.m. and its minimum at 7 p.m.

The daily variation is not constant, that is, it does not go through the same course on different days. The magnitude of the change is shown in Fig. 32, taken from the results of Dr.

Chree,[1] in which the curve O represents the variation in declination on ordinary days, Q that for very quiet days, and D that

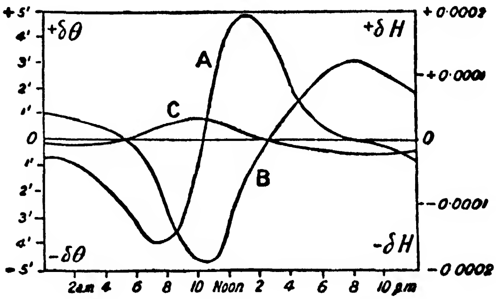


FIG. 31.

for days of considerable magnetic disturbance, the values taken being means over an eleven-year period.

Sir Arthur Schuster [2] has investigated the phenomenon of the daily variation, and has come to the conclusion that it is due to causes external to the earth, probably to electric currents in the atmosphere; and the daily magnetic variations cause induced currents in the earth, which reduce the amplitude of the vertical and increase that of the horizontal component. The earth currents which would produce these magnetic effects are of such a character that they indicate that the earth is not a uniformly conducting sphere; the upper layers conduct better than the lower layers. And further, the observed daily variation is similar in character to that which would be produced by the motion of the atmosphere due to the tidal action of the sun and moon, or periodic variations of the barometer, provided that the atmosphere is in such a state that the smallest electromotive force will produce current.

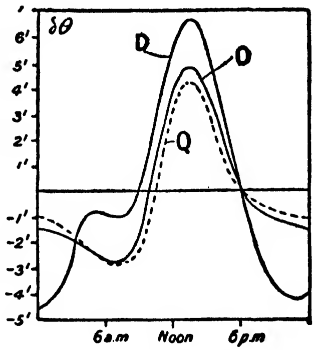


FIG. 32.

The daily variations in the horizontal component of the earth's field have been represented by v. Bezold in a very convenient form. A vector, representing in magnitude and direction the variation in H from the mean at any instant, is drawn from the point O (Fig. 33). During the day this vector makes a complete revolution, and its extremity describes the curve in the figure,

[1] C. Chree, *Phil. Trans.*, vol. 208, A. 1907.
[2] A. Schuster, *Phil. Trans. Roy. Soc.*, vol. 180, Part I. 1889.

upon which the appropriate times of day may be indicated, and the vector representing any required time may be seen. It is then found that for all points on the earth having the same latitude, the vector diagram of daily variation has the same shape, and by placing the axis AB in the direction of the meridian, the vector at any moment representing the variation can be added vectorially to that indicating the mean horizontal component of the earth's field at that point, to obtain the actual horizontal component and the direction in which the compass would set.

The daily variation is less in winter than in summer.

(iv) *Magnetic Storms.*—Simultaneous variations in the magnetic elements over the whole earth are frequently observed, the magnetic needles at the observatories undergoing rapid and sometimes enormous disturbances.

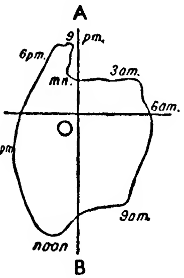


FIG. 33.

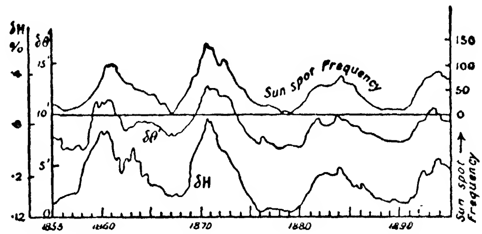


FIG. 34.

Eleven-Year Period.—On recording the frequency of sunspots and the magnitude of the daily variation in the magnetic elements, a surprising parallelism between the two phenomena becomes apparent (Fig. 34).

It is thus seen that the period of eleven years, during which the frequency of the occurrence of sunspots goes through a cycle, coincides with the period of change in the magnitude of the daily variations. In the diagram $\delta\theta$ is the amplitude of the daily variation in minutes of arc, and δH is the variation in the horizontal intensity expressed as a fraction of the whole amount. The diagram is given by A. Nippoldt,[1] and exhibits very clearly the parallelism in the three quantities.

Cause of Variations.—The origin of the earth's magnetism is still a matter for investigation, but the variations, although not thoroughly understood, have been explained on the assumption that the sun is emitting a radiation similar to the cathode rays, met with in a vacuum tube in which an electric current is passing (Chap. XIV). Such rays consist of minute particles, which, in passing through a gas such as our atmosphere, render it con-

[1] A. Nippoldt, *Erdmagnetismus, Erdstrom und Polarlicht.*

ducting for the electric current. Any potential difference between different localities in the upper layers of the atmosphere would then give rise to electric currents, and such currents having a magnetic field associated with them would, of course, affect the magnetic needle. Such phenomena as the eleven-year recurrence of the maximum daily variation lend colour to such a theory, for the radiation of all kinds from the sun changes with the nature of its surface. Also the daily variation itself may be due to the same cause; although it must not be forgotten that the changes in temperature due to the alteration in the amount of radiant heat received from the sun may help to cause the observed variations. It may also be noted that Bauer [1] noticed a small wave-like disturbance of the suspended needle, as the moon passed over the sun's disc, in the total eclipse of 1900, which was similar in character to the solar-diurnal variation, but smaller. This change took place at all the observing stations as the moon's shadow passed over them.

The daily variation has been analysed into three parts, a regular variation due to the sun, a minute variation due to the moon and an irregular disturbance due to the sun. The first of these is seen in the quiet day variation (p. 41) and is the largest of the three. It has been found possible to formulate currents in the conducting layers of the atmosphere which would produce the observed variation, the conductivity of the layers being influenced by radiation from the sun. The origin of the electromotive forces which produce the currents is not known definitely. The variation due to the moon is so small that it can only be found by the collection of observations over long periods. It is certainly connected with the phases of the moon.

Magnetic storms are not always accompanied by Auroral displays, but the latter are always associated with magnetic disturbances. Also the form of the Aurora Borealis is frequently such as might be explained by streams of cathode ray particles entering the magnetic field of the earth (Chap. XIV); and again, the spectrum of the Aurora is a line spectrum, showing that it is not reflected sunlight; and in it the lines of nitrogen, argon, neon and xenon have been detected, a fact which points to the conclusion that the light is emitted by the passage of an electric discharge through the atmosphere.

Recording Instruments.—For the purposes of a magnetic survey, portable instruments are necessary, since the magnetic elements at a great many places must be determined. These instruments (pp. 28–35) are not of great precision; and are incapable of measuring small variations in the earth's field. Consequently, at certain observatories, recording instruments or

[1] L. A. Bauer, *Terr. Mag. and Atmos. Elect.*, XV, 2. June, 1910.

magnetographs are erected, which give a permanent record of the small variations in the terrestrial magnetic field. The three types of instrument record respectively variations in declination, horizontal intensity and vertical intensity.

The *declination magnetograph* is an instrument in which a beam of light, reflected from a mirror attached to a suspended magnet, falls upon a sheet of photographically sensitive paper wound upon a drum which rotates at constant speed, the curve traced upon it indicating the variations in the declination. The instrument is thus a combination of the reflecting magnetometer and the chronograph.

In the instrument designed by Watson,[1] nine small permanent magnets are cemented in an aluminium centrepiece, which is suspended by a phosphor-bronze strip, the magnetic system being situated inside a massive block of copper, to cause oscillations to be rapidly damped out. The employment of phosphor-bronze for the suspension renders the reading of the instrument almost independent of temperature, for the elasticity of the phosphor-bronze decreases as the temperature rises, and the magnetic moment of the magnets likewise decreases. Hence the deflecting and the controlling couples rise or fall together, with alteration of temperature.

In the case of the *horizontal variometer* the suspended magnetic system is rotated, either by twisting the suspension or by means of compensating magnets, until its magnetic axis is perpendicular to the magnetic meridian. Eschenhagen,[2] using a quartz suspension, twisted the torsion head until the magnetic axis was 90° from the meridian. He also used two mirrors upon the suspended system, so that the doubling of the deflection due to reflection occurred twice, with corresponding increase in sensitiveness. If M be the magnetic moment of the system, θ the angle made with the meridian, and α the number of degrees of twist in the suspension,

$$\mathrm{MH}\sin\theta = c\alpha.$$

When, $\theta = 90°$, $\mathrm{MH} = c\alpha$, and $\mathrm{M}\delta\mathrm{H} = c\delta\alpha$.

$$\therefore \frac{\delta\mathrm{H}}{\mathrm{H}} = \frac{\delta\alpha}{\alpha}, \text{ and } \delta\mathrm{H} = \frac{c\delta\alpha}{\mathrm{M}}.$$

Thus, a given change δH will cause a bigger deflection, the smaller c and the greater the value of M. Hence a fine suspension fibre is used, and α is great, amounting to several revolutions, in order to maintain the magnet at right angles to the meridian.

As H varies so does θ, and the movement of the reflected beam of light is recorded photographically. θ never differs much from 90°. The scale of the record is calibrated by causing a deflection

[1] W. Watson, *Terrestrial Magnetism*, vi. 1901.
[2] M. Eschenhagen, *Terrestrial Magnetism*, v. 1900.

by means of a small magnet of known moment, placed at a distance from the instrument (see p. 6).

The *vertical intensity magnetograph* is usually a magnet, mounted as in the case of the dip circle, to rotate about a horizontal axis, the plane of rotation being the magnetic meridian. The end of the needle, which would ordinarily set upwards, is loaded until the needle is horizontal, and the horizontal component of the earth's field does not give rise to a couple tending to rotate the needle. Owing to the direction of the needle being perpendicular to the earth's vertical component, any variation in this causes corresponding variation in the position of equilibrium of the needle, and its movements are recorded by the beam of light and photographic drum as in the last two cases.

In order to respond to rapid changes in the magnetic field, the magnet must be as light as possible, and when supported upon knife-edges, any mechanical disturbance will cause such a needle to move about in azimuth, with loss in definiteness of the record ; also change in temperature produces alteration in the magnetic moment of the magnet, with resulting change in the position of equilibrium.

To get over these difficulties Watson [1] attached the magnets which are 8 cm. long and 1 mm. in diameter, to a quartz plate, to which the horizontal suspension fibres are fused. The arrangement is shown diagrammatically in Fig. 35. E is a slab of fused quartz, the upper face of which is polished and constitutes the mirror. The rods B and C are part of this, and serve both to carry the magnets M, M′, and for the attachment of the quartz fibres AB and CD. At A is a spring of fused quartz, to which is fused one end of the fibre AB. The attachments at B, C and D are all made by fusing the fibres on to the quartz rods, so that the suspension consists entirely of homogeneous fused quartz. At D is a torsion head, the adjustment of which serves to set the magnets horizontal. P is a 45° reflecting prism, to enable the readings to be made by means of a horizontal beam of light.

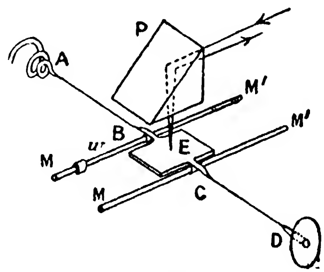


Fig. 35.

The small adjustable weight w is placed in such a position that the ends M of the needles, which usually point upwards, are now depressed below the level of the axis, and the magnets are brought into a horizontal position by rotating the torsion head D in a clockwise direction. The earth's vertical magnetic field tends to raise MM and depress M′M′, and any variation of the field is

[1] W. Watson, *Proc. Phys. Soc. Lond.*, XIX, II. 1904.

observed by the corresponding rotation of the magnets and the attached mirror E.

By the construction chosen, the apparatus is practically free from error due to change of temperature, for if this rises, the magnetic moment of the magnets decreases and the ends MM would be depressed. But the rigidity of the quartz fibres increases with rise of temperature, and so the couple exerted by them increases, causing the ends MM to be raised. The two effects are therefore opposite, and by adjusting the position of the small weight w, and the torsion in the fibres, the apparatus may be compensated for temperature change.

Geomagnetism.—The study of the propagation of seismic waves, generated by earthquakes, has revealed that inside the thick solid mantle of the earth is a core, of radius some 0·55 of the external radius of the earth, which does not transmit transverse waves involving shearing of the medium, and which must therefore be fluid. There is good reason to suppose that this core may have a finite electrical conductivity and may therefore carry electric currents.

The main magnetic field of the earth is explicable if certain plausible assumptions are made about these currents, which may be perpetuated by rotation in a field which the currents themselves assist to generate. Convection currents in the core may also play an important part. The energy to maintain the current system may thus come in part from the rotational energy of the earth and in part from heat, perhaps from radioactive minerals within the earth (p. 534).

On this theory, the various known large anomalies could be due to large eddies in the fluid core, and a small difference in the speeds of rotation of core and mantle could account for the slow westerly drift of the terrestrial field which constitutes the secular variation (p. 40). Much of this is still speculative, however.[1]

Local anomalies in the magnetic elements can often be correlated with non-uniformity in the geological structure and they are studied in the field, often in conjunction with gravity surveys and other physical investigations, in normal geological survey work and in prospecting for oil and other minerals.

Many sedimentary deposits, such as some clays, are weakly magnetised, and the direction and intensity of magnetisation in the various layers may be attributed to changes in the orienting effect of the terrestrial magnetic field when the particles were settling. In this way some idea of the past magnetic history of the earth is being built up.[2]

[1] See S. K. Runcorn, *Science Progress*, XXXVIII, p. 668. 1950. Also S. Chapman, " The Earth's Magnetism." Methuen & Co., 2nd edition, 1951.

[2] S. K. Runcorn, *loc. cit.*

CHAPTER II

THE ELECTRIC CURRENT

Under certain circumstances an ordinary metallic wire may exhibit distinctive phenomena, the most striking being, the existence of a magnetic field in the space surrounding it, and the production of heat in it. We say, then, that an electric current is flowing in it. The electric current is always accompanied by these two effects, and their presence may be taken to indicate its existence. Our reason for speaking of this phenomenon as a current, which implies a flow of something along the wire, rather than as a statical condition, will appear later.

Magnetic Field accompanying a Current.—An infinitely long straight wire, carrying an electric current, is surrounded by circular lines of magnetic force, the centres of the circles lying upon the axis of the wire, their planes being perpendicular to it.

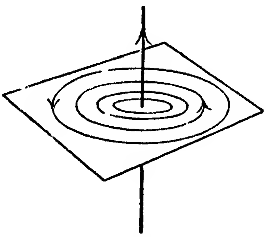


Fig. 36.

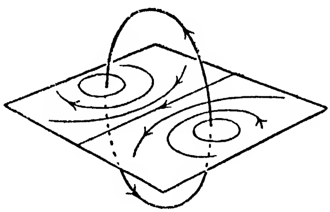


Fig. 37.

In the case of any straight piece of wire which is fairly long, the magnetic lines may easily be mapped out, either by the method of iron filings, or with a small compass needle, and will be found to be approximately circular (Fig. 36). In the case of a circle of wire carrying a current (Fig. 37), the magnetic lines of force are not so simple in shape as in the case of a straight wire, but they may be found in a similar manner, and it will be noticed that for a small region near the centre of the circle, the field is nearly uniform, that is, the lines are nearly parallel. The fundamental experiments exhibiting the presence of a magnetic field when an electric current is flowing are due to Oersted (1820), the existence of the current having been recognised by certain other effects for the previous twenty years.

Ampère's Theorem.—In his celebrated memoir[1] of 1823, Ampère stated that "Every linear conductor carrying a current is equivalent to a simple magnetic shell, the bounding edge of which coincides with the conductor, and the moment of which per unit of area, that is, the strength of the shell, is proportional to the strength of the current." By a magnetic shell is meant an infinitely thin sheet of material, magnetised in a direction at right angles to the surface of the sheet, so that one side of the sheet is a N, and the other a S, polar surface. The form of the magnetic field due to a current may therefore be calculated by means of purely magnetic considerations, on replacing the current circuit by its equivalent magnetic shell; but the same result may also be obtained for a complete circuit by treating each small element of it as a straight current of length δl, and applying the relation $H \propto \frac{i\delta l \sin\theta}{r^2}$, where H is the resulting strength of magnetic field, i the current, r the distance from the element of the circuit to the point at which H is to be found, and θ the angle between the direction of the current and the line joining the element to the point. The direction of the field is at right angles to the plane containing the element and the line joining it to the point, and is indicated in Fig. 38.

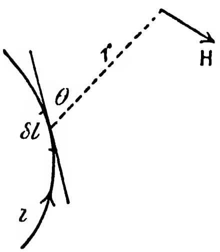


FIG. 38.

This law was proved by Biot and Savart to hold in the case of a long straight wire carrying current. The magnetic fields, as determined by this method and by that of the equivalent magnet shell, are identical. It may be noted that the method of the magnetic shell is the more satisfactory, as, however difficult the problem may be, it is always possible to find a magnetic shell, or system of shells, that will be equivalent to the current; but the method of the circuit element previously considered fictitious, may be justified, as we now know that the current is the motion of electric particles (Chap. XVI).

Unit of Current.—The magnetic field due to a current being the most constant and the simplest of the accompanying phenomena, it is chosen for the purpose of measuring the current. The unit strength of magnetic field being established (p. 3), we may now define our unit of electrical current in terms of it.

Thus, *the unit current is one that is equivalent to a magnetic shell of unit strength*; or, by means of the relation above, we may define the unit of current as that which will enable us to replace the sign of variation by one of equality, so that

[1] *Théorie des phénomènes électro-dynamique, Mémoires de l'Institut,* IV. 1823.

$$H=\frac{i\delta l \sin \theta}{r^2}.$$

In this equation δl is a very small quantity, and hence if r is to be constant, while δl is increased to finite size, the circuit must evidently be in the form of a circle. Thus the field at the centre of the circle is $\frac{il}{r^2}$ and is at right angles to its plane, l being the total length of arc in which the current flows. If then i, l and r are all unity, the field is of unit strength, and we have the ordinary definition of unit current, as *that current which flowing in an arc of a circle of unit length, the radius being unity, produces unit magnetic field at the centre.* If the circle consist of n complete turns $l=2\pi nr$, and it follows that

$$H=\frac{2\pi ni}{r}.$$

Tangent Galvanometer.—The last equation is employed in a form of instrument for measuring an electric current in terms of a magnetic field and a deflection, and from the form of the relation between the current and the deflection, the name "tangent galvanometer" is given to the instrument. It is essentially a magnetometer in which the magnetic field, F, is due to the current flowing in a vertical circular coil, whose plane is in the magnetic meridian. The two fields in which the needle is situated are therefore F, due to the coil, and H, the horizontal component of the earth's field (Fig. 39); and the needle is in equilibrium when its magnetic axis makes an angle $\theta=\tan^{-1}\frac{F}{H}$ with the meridian.

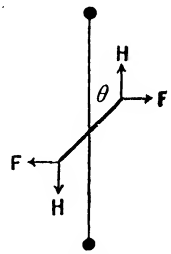


FIG. 39.

Then since—

$$F=\frac{2\pi ni}{r}$$

$$\frac{2\pi ni}{rH}=\tan \theta, \text{ or, } i=\frac{rH}{2\pi n} \tan \theta.$$

A common type of the apparatus is shown in Fig. 40, two coils being provided, one of 2 turns of thick wire for use with large currents, and one of about 20 turns for use with smaller currents. The deflection is observed by means of a fine pointer which moves over a horizontal circular scale, parallax being avoided by placing the eye vertically over the needle in taking a reading, the eye being moved until the image of the needle in a plane mirror lying underneath it appears to coincide with the needle itself.

In principle the tangent galvanometer resembles the magnetometer described on p. 7, the linear scale and bar magnet being replaced by the circular coil carrying the current.

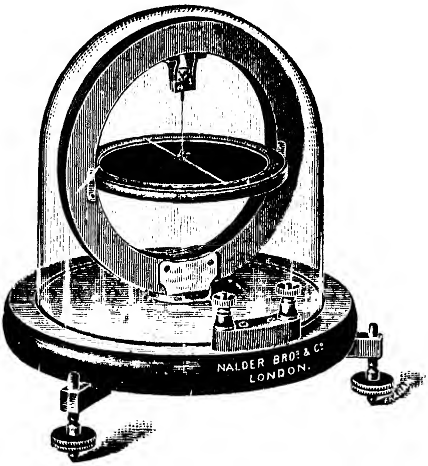


FIG. 40.

The same precautions with respect to reading both ends of the pointer, and reversing the deflection so that the reading is made on the other side of zero, are made, as in the case of the magnetometer, and further the pointer must be adjusted to be at right angles to the needle. Owing to the fact that the magnetic field at the centre of the coil is sensibly uniform for only a small area, the needle must be as small as possible; if too large it is not in a uniform field, and the relation given above will not hold good.

In making observations, the deflections should be neither too great nor too small. If too small, any error in reading is a very large proportion of the whole deflection, and if too great, the tangent of the deflection increases so rapidly that a small change in deflection means a very large actual change in the value of the current. To find the position on the scale for readings of greatest accuracy, consider $\delta\theta$ to be a small increment in the deflection corresponding to the increment δi in the current. $\frac{\delta i}{i}$ is the relative change in the current, and $\frac{\delta i}{100 i}$ the percentage change in the current corresponding to $\delta\theta$, and for greatest accuracy $\frac{di}{i}$ should therefore be as small as possible.

Since the current is proportional to the tangent of the deflection, $i = \text{K} \tan\theta$,

$$\text{and, } \frac{\delta i}{\delta\theta} = \text{K} \sec^2\theta, \text{ when } \delta i \text{ and } \delta\theta \text{ are infinitesimal};$$

$$\therefore \frac{\delta i}{i} = \frac{\sec^2\theta}{\tan\theta} \cdot \delta\theta = \frac{2}{\sin 2\theta}\delta\theta.$$

Hence for $\frac{\delta i}{i}$ to be as small as possible for a given value of

$\delta\theta, \frac{2}{\sin 2\theta}$ must be as small as possible; *i.e.* $\sin 2\theta$ must be as great as possible. This occurs when $2\theta = 90°$, or $\theta = 45°$. Hence it is desirable to make the deflection as near to 45° as possible, whenever accuracy in determining the current is required. The curve, Fig. 41, shows the relative accuracy in determining the current when the deflection varies from 0° to 90°, taking the accuracy at 45° as 100. It will be seen that for the accuracy not to fall to 50 per cent. of the maximum, the deflection must lie between 15° and 75°.

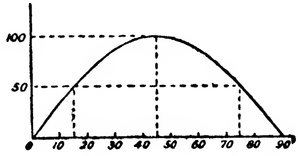


FIG. 41.

A more refined instrument is shown in Fig. 42. In this case there are two coils, the distance between them being equal to the radius of either coil, the object being to obtain the most uniform field in which to suspend the needle.

In order to find the strength of magnetic field at a point on the axis of a circular coil at a distance x from the centre, consider

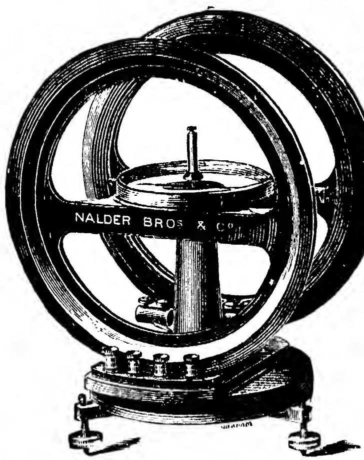


FIG. 42.

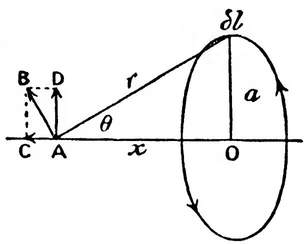


FIG. 43.

the field due to an element δl of the circle. The field at A (Fig. 43) due to this element is $\frac{i \cdot \delta l}{r^2}$, where i is the current in the coil, and is in the direction AB, at right angles to the plane containing r and δl. This may be resolved into two components, AC along the axis, and AD at right angles to it. Every element of the circular coil will produce a component at right angles to the axis and in each case it is parallel to the radius of the circle drawn to the element. Hence, taking the whole circle, these

components of the field corresponding to AD will give a resultant zero. The components corresponding to AC, which are along the axis, will on the contrary be added together.

Component AC due to element δl

$$=\frac{i\,.\,\delta l}{r^2}\cdot\frac{\mathrm{AC}}{\mathrm{AB}}=\frac{i\,.\,\delta l}{r^2}\cdot\frac{a}{r}=\frac{ia\delta l}{r^3}.$$

For the whole circle, δl must be replaced by $2\pi a$.

$$\therefore\ \text{Field}=\frac{2\pi a^2 i}{r^3},$$

and if the coil consist of n turns sufficiently close together to take a mean radius without introducing a sensible error,

$$\text{Field}=\frac{2\pi n a^2 i}{r^3}$$

Again, $r^2=x^2+a^2$;

$$\therefore\ \text{Field}=\frac{2\pi n a^2 i}{(x^2+a^2)^{\frac{3}{2}}}$$

The strength of field at a point upon the axis is therefore greatest at the centre of the circle, for here $x=0$ and the expression becomes $\frac{2\pi n i}{a}$. It decreases as we pass away from the centre, becoming zero at infinity, but its rate of change as we pass away from the centre is not constant. The rate of change from point to point along the axis is the differential coefficient of the above expression with respect to x, that is $\frac{d}{dx}\left[\frac{2\pi n a^2 i}{(x^2+a^2)^{\frac{3}{2}}}\right]$, and it is of importance to find whether there is any point upon the axis at which this *rate of change* becomes constant. Calling it y we see that if it is constant, then

$$\frac{dy}{dx}=0,\ \text{or,}\ \frac{d^2}{dx^2}\left[\frac{2\pi n a^2 i}{(x^2+a^2)^{\frac{3}{2}}}\right]=0.$$

Now, $2\pi n a^2 i$ is constant, so that in dealing with rates of change it may be omitted; also remembering that

$\frac{1}{(x^2+a^2)^{\frac{3}{2}}}=(x^2+a^2)^{-\frac{3}{2}}$, we have—

$$\frac{d}{dx}(x^2+a^2)^{-\frac{3}{2}}=-3x(x^2+a^2)^{-\frac{5}{2}}$$

$$\frac{d^2}{dx^2}(x^2+a^2)^{-\frac{3}{2}}=-3\{(x^2+a^2)^{-\frac{5}{2}}-5x^2(x^2+a^2)^{-\frac{7}{2}}\}$$

Putting this equal to zero and dividing throughout by $-3(x^2+a^2)^{-\frac{5}{2}}$, we have—

$$5x^2(x^2+a^2)^{-1}=1,$$

$$\therefore\ 5x^2=x^2+a^2,\ 4x^2=a^2,\ \text{or}\ x=\frac{a}{2}.$$

Thus, at the point on the axis whose distance from the plane of the circle is $\frac{a}{2}$, the rate of change of the field as we pass along the axis becomes constant.

This fact is made use of in the Helmholtz pattern of galvanometer (Fig. 42); the two coils are placed coaxially and at a distance apart equal to the radius of either, the rate of change of field being most uniform at a point midway between them, which point is at a distance $\frac{a}{2}$ from each coil. Any diminution in field due to one coil as we pass away from this point is compensated for by the equal increase in the field due to the other coil, the rate of change being here constant and occurring in opposite directions for the two coils. Substituting the value $\frac{a}{2}$ for x in the expression for the field on the axis of a coil, and remembering that there are two coils, we have

$$F=\frac{4\pi na^2i}{\left(\frac{5a^2}{4}\right)^{\frac{3}{2}}}=\frac{32}{\sqrt{5^3}}\cdot\frac{\pi ni}{a}.$$

And the expression for the deflection is

$$\frac{32}{\sqrt{5^3}}\cdot\frac{\pi ni}{aH}=\tan\theta,$$

$$\text{or, } i=\frac{\sqrt{125}aH}{32\pi n}\tan\theta.$$

The above reasoning may be illustrated by plotting the values of the field strength due to a coil at different distances from the centre. The curve PQR (Fig. 44) is obtained, which will be seen to be straight at the point Q at a distance $\frac{a}{2}$ from the coil, since the curve changes here from being convex upwards to being concave upwards. The rate of change of field at this point is constant, that is, its variation is zero. The curve SQT is plotted for the second coil, and the curve M is obtained by adding the ordinates of the other two, and represents the resultant field due to both coils. It is easily seen that for some distance on either

side of the point midway between the coils the field is fairly constant, the reason being that the two curves are straight at Q, so that the falling off in either direction of one of the fields is balanced by the equal increase of the other field.

In Fig. 45, the lines of force for the double coil are drawn, and it will be seen that in the middle of the field there is a region of

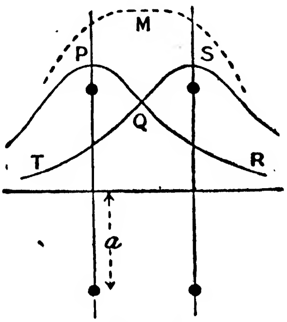


FIG. 44.

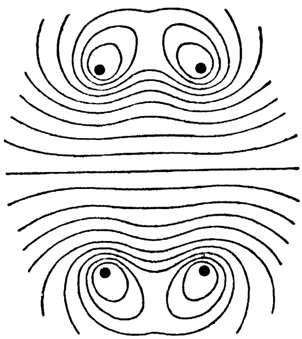


FIG. 45.

considerable extent where the magnetic field is approximately uniform.

Electromotive Force and Potential Difference.—Whenever a current flows in a conductor, heat is developed in it, the amount of heat developed being proportional to the time for which the current flows; moreover, it is always found necessary to apply some external agency to maintain a steady current. This implies that work is being expended in order to maintain the current, the energy being drawn from the outside source. The rate of expenditure of energy required to maintain the current is measurable in terms of the rate of production of heat, when there is no change taking place in the conductor. From analogy with the flow of water in a closed circuit, in which case the flow is maintained by some mechanical agency such as a pump at some point or points, we consider that the electric current is maintained by an *electromotive force.* Just as we took the magnetic field in the neighbourhood of a current as a measure of the current itself, so we may take the rate at which energy is expended to maintain a current to measure the necessary electromotive force. Then for a given current, the electromotive force (or more shortly the E.M.F.) required to maintain it, is proportional to the rate at which energy is expended, and by choosing suitable units we arrive at our unit of E.M.F. Thus, *when the current is unity and the rate of working is one erg per second, the E.M.F. is unity,* and in terms of these units

$$\text{Rate of working} = (\text{current} \times \text{E.M.F.}) \text{ ergs per second.}$$

Thus, if an E.M.F. e maintain a current i, for t seconds,

$$\text{Work done} = eit \text{ ergs.}$$

It must be remembered that a given conductor is usually only part of the circuit, and somewhere in the circuit is the source of the energy required to maintain the current. This source may be an electric battery, in which case the ultimate source of the energy may be some chemical reaction occurring in the battery ; or it may be a dynamo-electric machine, in which case the energy is derived from some kind of heat engine, or it may be one of a number of other sources ; but in any case the rate of working to maintain unit current in the circuit is called the electromotive force in the circuit. The resulting heat may be liberated in various parts of the circuit, and this will take place according to laws which we must now examine.

An electromotive force always acts in one direction in the circuit, and if there be a number of electromotive forces in the same circuit, the excess of those acting in one direction over those acting in the other direction is the resultant or effective electromotive force in the circuit. Thus an electromotive force is a directed quantity and in this respect is analogous to a mechanical force. In many mechanical processes the energy supplied by the driving force is eventually dissipated as heat, and the rate at which the heat is developed at various points depends upon the frictional resistance to motion at these points. Similarly, in the case of an electrical circuit the energy supplied by the source of E.M.F. appears as heat in the circuit, but the rate of production of heat at any point, for any given current, varies according to the nature of the conductor at that point.

For a given conductor, the work converted into heat in it in one second when unit current flows is called the *potential difference* between its ends. Thus, potential difference is measured in the same units as electromotive force, but they have this difference, that an electromotive force has always the same direction in the circuit, whereas the potential difference has a direction depending on that of the current. If the current flows in a circuit in the direction in which the electromotive force tends to produce current, the source of electromotive force transfers energy to the circuit and the energy of the source decreases. But if the direction of the current is reversed so that the electromotive force opposes its flow, the energy of the source of electromotive force increases at the expense of the energy of the source of greater electromotive force which is maintaining the current in the circuit. On the other hand, a potential difference always corresponds to the dissipation of energy in form of heat in the circuit, whichever way the current

flows. Thus, if electromotive force corresponds to a motive mechanical force, potential difference corresponds to a frictional force, which depends for its direction upon the direction of motion, and is a measure of the heat produced per second between two points, for a given continuous motion of matter between one point and the other.

Ohm's Law.—In the chapter on Electrostatics (IV) we shall see that potential difference may be measured quite independently of any current flowing, and if for any conductor the potential difference be measured by this independent means, and the current also be measured, say, by the tangent galvanometer, it will be found that for the case of an ordinary metallic conductor there is a simple relation between potential difference and current, the current is proportional to the potential difference. This relation was first clearly stated by G. S. Ohm [1] and is known as Ohm's law. Although Ohm had not the means of establishing the law with any great certainty, later experimenters verified it to a high degree of accuracy.

Resistance.—Ohm's law may be expressed in the form,

$$\frac{\text{potential difference}}{\text{current}}=\text{constant},$$

for any conductor under constant physical conditions. The name *resistance* has been given to this constant, and the name *conductance* to the inverse of it.

$$\text{Thus, } \frac{p.d.}{i}=\text{resistance}, \quad \frac{i}{p.d.}=\text{conductance}.$$

The unit of resistance follows at once from the units of potential difference and of current, and is the resistance of a conductor in which unit potential difference corresponds to unit current, or the potential difference between the ends of the conductor is unity when unit current flows in it.

$$\text{Thus, } \frac{p.d.}{i}=r, \text{ or, p.d.}=ir.$$

$$\text{And, rate of working}=p.d.\times i=i^2r=\frac{(p.d.)^2}{r} \text{ ergs per second.}$$

Practical Units.—Although the centimetre, gramme and second are of convenient size for the measurement of length, mass and time, the derived units of electromotive force, current and resistance, resulting from them, and called the absolute C.G.S. units, are not of convenient size for ordinary electrical purposes. Hence a new unit of current is chosen which shall be simply

[1] G. S. Ohm, *Die galvanische Kette mathematisch gearbeitet.* Berlin, 1827.

related to the old unit, but of more useful size; it is called the *ampere*, being named after the celebrated experimenter Ampère, and is one-tenth of the size of the absolute unit. Thus the expression for the field at the centre of a circular coil will be $\frac{2\pi n I}{10r}$, where I is the current in amperes.

Similarly the unit of E.M.F. is chosen to be of the order of that of an ordinary electric cell; it is 100,000,000 or 10^8 absolute units, and is called the *volt*.

Thus for a complete circuit, rate of working

$$\begin{aligned} &= ei \text{ ergs per second} \\ &= E \,.\, 10^8 \times I \,.\, 10^{-1} \\ &= EI \times 10^7 \text{ ergs per second,} \end{aligned}$$

where E is now measured in volts, and I in amperes. Upon this system the practical unit of rate of working is the work done per second when one volt maintains a current of one ampere, and is called the *Watt*. We see then that rate of working = EI watts, and further that one watt $= 10^7$ ergs per second. The engineer's unit of rate of working, the horse-power, or 33,000 foot-pounds per minute, may be converted into watts by converting feet to centimetres, pounds weight to dynes, and minutes to seconds, when it will be found that one horse-power = 746 watts (approx.).

Again, the name *Joule* is given to the unit of work upon the practical system; it is the work done in one second when a current of one ampere is maintained by an E.M.F. of one volt, so that one watt = one joule per second.

The heat developed in any circuit by an electric current may therefore be found in terms of the current flowing in it and the electromotive force which maintains the current, when the mechanical equivalent of one calorie, or as it is termed, Joule's equivalent, is known. This quantity has been determined in a number of ways, but the mean value may be taken to be $4{\cdot}18 \times 10^7$. Thus the conversion of $4{\cdot}18 \times 10^7$ ergs into heat would raise one gramme of water one degree Centigrade. Then the rate of working may be expressed either in ergs per second or calories per second. Thus,

$$\text{Rate of working} = EI \times 10^7 \text{ ergs per second,}$$

$$= \frac{EI \times 10^7}{4{\cdot}18 \times 10^7} = EI \times 0{\cdot}239 \text{ calories per second.}$$

The practical unit of resistance is called the *ohm*, and is the resistance of a conductor for which there is a potential difference of one volt between its ends when a current of one ampere flows

in it. Calling R the resistance of any conductor in terms of this unit, we see that

$$\frac{E}{I}=R, \text{ or, } E=IR.$$

$$\text{Again, one ohm}=\frac{\text{one volt}}{\text{one ampere}}=\frac{10^8 \text{ absolute units of p.d.}}{10^{-1} \text{ absolute unit of current.}}$$

$\therefore$ One ohm is equal to 10^9 absolute units of resistance.

$$\text{Also, rate of working}=EI=I^2R=\frac{E^2}{R} \text{ watts,}$$

$$=EI\times 0{\cdot}239 \text{ calories per second,}$$

from which we see that for a given conductor, the heat produced in it per second is proportional to the square of the current.

Combination of Resistances.—From the definition of resistance given above, we may find the resistance of a number of conductors combined in series, that is, end to end, so that the same current flows through all of them ; or in parallel, in which case they are all joined side by side between two points so that the current is divided between them.

FIG. 46.

(i) *Series.*—The current I enters at the point A (Fig. 46), and flows through all the resistances, leaving at D. Then if there is no source of E.M.F. in any of the conductors,

p.d. between A and B $=IR_1$
,, ,, B and C $=IR_2$
,, ,, C and D $=IR_3$

$\therefore$ p.d. between A and D $=IR=IR_1+IR_2+IR_3$

where R is the effective resistance between A and D.

$$\therefore R=R_1+R_2+R_3+ \ldots$$

That is, the combined resistance is the sum of the separate resistances.

(ii) *Parallel or Multiple Arc.*—In this arrangement of conductors the current enters at the point A (Fig. 47), and divides into a number of parts which unite again at B. Taking the total current I, equal to the sum of the separate currents,

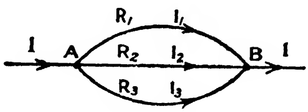


FIG. 47.

$$I=I_1+I_2+I_3.$$

If, now, the p.d. between A and B is equal to E,

$$E=IR=I_1R_1=I_2R_2=I_3R_3$$

where R is the effective resistance between A and B.

$$\therefore\ I=\frac{E}{R},\ I_1=\frac{E}{R_1},\ I_2=\frac{E}{R_2},\ I_3=\frac{E}{R_3}.$$

$$\text{Hence, } \frac{E}{R}=\frac{E}{R_1}+\frac{E}{R_2}+\frac{E}{R_3}.$$

Dividing by the common quantity E, and writing the expression for any number of conductors, we have—

$$\frac{1}{R}=\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}+\frac{1}{R_4}+\ \ldots$$

Thus for conductors in parallel the combined conductance is the sum of the separate conductances, the conductance being defined as the reciprocal of the resistance.

To find the current in any one branch, we may note that

$$I_1=\frac{E}{R_1}=\frac{IR}{R_1},$$

i.e. current in one branch $=$ main current $\times\dfrac{\text{combined resistance}}{\text{resistance of branch}}$.

Resistivity or Specific Resistance.—The resistance of a conductor depends upon its dimensions, and also upon the material of which it is made. The resistance of a conductor of unit length and unit area of cross section is called its *resistivity* or *specific resistance*, S, and the inverse of this is its *conductivity*. The shape of the cross section is immaterial. Knowing the resistivity of the material we may readily find the resistance of a uniform conductor of any dimensions. For the conductor may be considered to consist of a number of unit conductors S, so that if these are placed end to end we have a number l of these in series, where l is the length of the conductor. Since these unit conductors are in series, the resistance of this rod of unit section is Sl. If, now, a of these are situated in parallel, a is the total area of cross section, and the combined resistance is given by

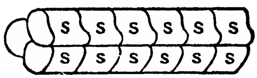


FIG. 48

$$\frac{1}{R}=\frac{1}{Sl}+\frac{1}{Sl}+\frac{1}{Sl}+\ \ldots \text{ to } a \text{ terms}$$

$$=\frac{a}{Sl}.$$

$$\therefore\ R=\frac{Sl}{a},\ \text{or, } S=\frac{Ra}{l}.$$

We can therefore find the resistance R if the resistivity S is known, and *vice versâ.* The universal method for finding the resistivity is to measure the resistance of the conductor by one of the methods described in Chapter III, and, knowing its dimensions, to calculate the resistivity. The resistivity of a number of substances is given in the second column of the Table below, in which the data are taken from Kaye and Laby's Tables, the resistivity being in International Ohms per unit conductor.

Substance.	Resistivity.	Temperature.	Temperature coefficient of resistivity.
Aluminium . .	$3{\cdot}21 \times 10^{-6}$ (Jaeger & Diesselhorst)	18° C.	0·0038
Copper . . .	$1{\cdot}59 \times 10^{-6}$ (Mean of number)	18° C.	0·00428
Gold	$2{\cdot}42 \times 10^{-6}$ (J. & D.)	18° C.	0·0040
Iron (soft) . .	$13{\cdot}9 \times 10^{-6}$ (J. & D.)	18° C.	0·0062
Lead . . .	$2{\cdot}08 \times 10^{-5}$ (J. & D.)	18° C.	0·0043
Mercury . .	$9{\cdot}407 \times 10^{-5}$ (International convention) 0° C.	0° C.	0·0009
Platinum . .	$1{\cdot}10 \times 10^{-5}$ (J. & D.)	18° C.	0·0038
Silver . . .	$1{\cdot}63 \times 10^{-6}$ (J. & D.)	18° C.	0·0040
Manganin (Cu 84, Ni 4, Mn 12)	$4{\cdot}205 \times 10^{-5}$ (J. & D.)	18° C.	0·000025
Platinoid (Cu 62, Ni 15, Zn 22)	$3{\cdot}44 \times 10^{-5}$ (Lees) 18° C.	18° C.	0·00025
Constantan (Cu 60, Ni 40)	$4{\cdot}9 \times 10^{-5}$ (J. & D.)	18° C.	−0·000002 to +0·00001
Nichrome . .	$1{\cdot}10 \times 10^{-4}$ (N.P.L.)	20° C.	0·00017

Electrolysis.—In the early years of the nineteenth century it was found that when an electric current flows in a solid or liquid substance which is not a metal, chemical action occurs, the products of the chemical action appearing at the conductors by which the current enters or leaves the substance. Non-metallic substances will not, as a rule, carry a current, unless some such chemical action occurs. To Faraday we owe the quantitative account of the phenomenon and the nomenclature now universally applied to it. The substance carrying the current, which in the act of carrying it undergoes decomposition, is called an *Electrolyte*, the conductors by which the current enters and leaves the electrolyte are called *Electrodes*, that at which the current enters being the *Anode*, and that at which it leaves, the *Cathode*, while the name *Electrolysis* is given to the whole process. The metals are liberated at the cathode, and the acid radicles at the anode; but owing to secondary reactions at the electrodes it often happens that the substance is dissolved in the electrolyte, or

combines with the electrode; in that case it does not appear in the free state.

The most common electrolytes are solutions of inorganic salts or acids in water. For example, if a current be passed through a dilute solution of sulphuric acid, platinum plates being used as electrodes, hydrogen appears at the cathode in the form of bubbles, and SO_4 is liberated at the anode,

$$H_2SO_2 = H_2 + SO_4.$$

This SO_4 does not appear in the free state, but with the water of the solution again forms sulphuric acid,

$$2H_2O + 2SO_4 = 2H_2SO_4 + O_2.$$

Bubbles of oxygen form at the anode, and if the gases at the cathode and anode be collected, it will be found that the hydrogen has twice the volume of the oxygen. If a copper anode had been employed, copper sulphate would have been formed in place of the oxygen.

Faraday's Laws.—As the result of Faraday's work, two laws were enunciated which bear his name.

(i) The amount of decomposition is proportional to the current and to the time for which it passes.

(ii) The amounts of different substances liberated by the same current, flowing for the same time, are proportional to the chemical equivalents of the substances.

From these two laws it follows that if the amount of any one substance liberated for a given current in a given time be known, the amount for any other substance may be found, provided that its chemical equivalent, which in the case of an element is the atomic weight divided by its valency, be known.

The mass of a substance liberated by one ampere in one second is called its *Electro-chemical Equivalent.*

The most accurately measured electro-chemical equivalent is that of silver, for which the most recent determination gives the value 0·0011183, and from this we can calculate that of any other substance. For example, the electro-chemical equivalent of di-valent copper is $0{\cdot}0011183 \times \frac{63{\cdot}6}{107{\cdot}9 \times 2} = 0{\cdot}000329$; the atomic weight of silver being 107·9, and that of copper 63·6.

The theory of electrolysis will be left to a later chapter, but we may note here that the constancy of Faraday's laws, and the ease with which the mass of the metallic substance liberated may be accurately determined, afford a ready means of measuring an electric current. The apparatus for carrying out the electrolytic measurements is called a *Voltameter*, and there are three types of voltameter frequently employed. In all cases the mass

of substance liberated in a known time is observed, and the current calculated from the relation

$$M = Izt,$$

where z is the electro-chemical equivalent.

		Atomic weight (O=16).	Valency.	Electro-chemical equivalent.
Aluminium . .	(Al)	26·97	3	
Antimony . .	(Sb)	121·76	3	
Bismuth . . .	(Bi)	209·0	3	
Bromine . . .	(Br)	79·92	1	0·0008284
Cadmium . . .	(Cd)	112·4	2	
Calcium . . .	(Ca)	40·08	2	
Chlorine . . .	(Cl)	35·46	1	0·0003676
Copper . . .	(Cu)	63·54	1 or 2	0·0003293
Gold	(Au)	197·2	3	
Hydrogen . . .	(H)	1·0080	1	0·00001045
Iodine	(I)	126·9	1	
Iron	(Fe)	55·85	2 or 3	
Lead	(Pb)	207·2	2	
Mercury . . .	(Hg)	200·6	1 or 2	
Oxygen . . .	(O)	16·00	2	0·00008293
Platinum . . .	(Pt)	195·2	4	
Potassium . .	(K)	39·096	1	
Silver	(Ag)	107·88	1	0·0011183
Sodium . . .	(Na)	22·997	1	0·0002384
Tin	(Sn)	118·7	2 or 4	
Zinc	(Zn)	65·38	2	0·0003387

Water Voltameter (Hoffmann's Tube).—Water slightly acidulated with sulphuric acid is employed as the electrolyte, the hydrogen liberated at the cathode (Fig. 49) being collected in the graduated tube, on the passage of the current for a known time. Allowance must be made for the fact that the gas is not under standard conditions. From the difference of levels of the liquid in the tube and the reservoir, the difference between the pressure of the hydrogen and the atmospheric pressure is known in terms of liquid column, and dividing by the density of mercury we obtain the amount to be added to the height of the barometer to give the actual pressure of the hydrogen. From this must be deducted the maximum vapour pressure of water at the temperature of the tube, which may be found from Regnault's tables, in order to obtain the pressure of the dry hydrogen. Calling this P, and the temperature $t°$ C., the volume reduced to 76 cm pressure and 0° C. is—

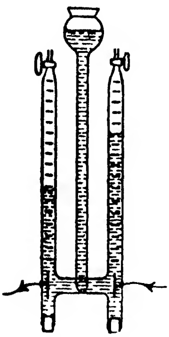


FIG. 49.

$$\frac{V \times P \times 273}{76 \times (t + 273)},$$

where V is the observed volume of hydrogen. The density of hydrogen at 0° C. and 76 cm. pressure being 0·08987 gm. per litre, the mass of hydrogen liberated is known, and the electro-chemical equivalent being 0·00001045, the current can be calculated.

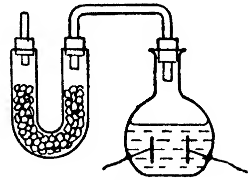


FIG. 50.

Another form of water voltameter is shown in Fig. 50, the hydrogen and oxygen liberated escaping together, and the water vapour carried away with them being caught by the drying tube. The whole apparatus is weighed before and after the passage of the current, the loss in weight being that of the water decomposed by the current. The mass of water decomposed by the passage of one ampere for one second is—

$$0{\cdot}00001045 \times \frac{(16+2{\cdot}016)}{2} = 0{\cdot}00001045 \times 9{\cdot}008,$$

0·00001045 being the electro-chemical equivalent of hydrogen.

For practical purposes this form of the apparatus is superior to the Hoffmann's tube, as the result depends upon weighing instead of upon measurement of volume, and further, the current may be passed for a much longer time, since in this case there is no question of the tube becoming filled with gas.

Copper Voltameter.—The cathode C (Fig. 51) consists of a thin copper sheet suspended from a stout conductor, and the anode of two sheets, one on either side of the cathode, so that the deposition of copper takes place on both sides of it. A and B are two wooden bars which carry the leads. They rest upon the edge of a jar or beaker containing a water solution of copper sulphate, made by dissolving copper sulphate crystals in about four times their weight of water, a few drops of concentrated sulphuric acid being added. The current employed should not be too great, or the copper deposited will be in a soft, friable condition, and will therefore be liable to be washed off the plate. Provided that the current does not exceed one ampere for each 50 square centimetres of cathode, the deposit will be hard, bright, metallic copper. The cathode must first be well cleaned with emery paper, then, when the current in the circuit has been adjusted to a suitable value, the cathode is removed from the cell, well washed, dried and weighed. It is then replaced in the cell and the current passed for a known time. The cathode

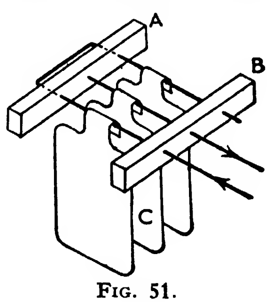


FIG. 51.

is then removed and again washed, dried and weighed. The increase in weight, divided by the time and by 0·0003293, the electro-chemical equivalent of copper, gives the value of the current.

In the process of electrolysis the copper sulphate is decomposed, the copper being liberated at the cathode, and SO_4 at the anode. The latter combines with the copper of the anode forming copper sulphate, so that the total amount of the copper sulphate in the solution remains unchanged. The loss in weight of the anode must not be taken as a measure of the current, since it includes not only the amount of copper which has gone into solution, but also any impurities which have become detached as the copper plate is dissolved.

This form of voltameter is very widely used for the calibration of ammeters and tangent galvanometers, as it is extremely simple in form and easily made, and by means of it the current may be determined to within an error of one part in several hundred.

Silver Voltameter.—When great accuracy is required, the silver voltameter of Lord Rayleigh's pattern may be employed (Fig. 52).

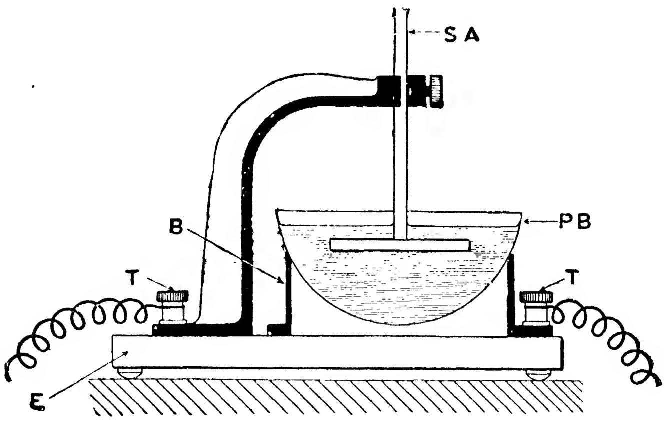


FIG. 52.

The cathode is a platinum basin, PB, and the anode a plate of pure silver, the electrolyte being a solution of 15 to 20 grammes of pure silver nitrate in 100 grammes of water. Metallic silver is deposited upon the platinum dish, and, owing to its high electro-chemical equivalent, the mass deposited for a given passage of current is greater than in the case of copper in the copper voltameter. The acid radicle NO_3 liberated at the anode by the process of electrolysis forms silver nitrate with the metal

of the anode itself, which is thereby dissolved. As this process of solution of the anode goes on, any impurities in the silver are liberated and these, together with the disintegrated silver, would fall upon the platinum plate if not prevented from doing so. To this end the anode is wrapped in a piece of pure filter paper, which, being permeated by the solution, will not prevent the passage of the current, but will catch the impurities. The current employed should not exceed 0·03 ampere per square centimetre of surface of cathode. The current is calculated from the deposit and the time, just as in the previous cases.

Kirchhoff's Laws.—We owe to Kirchhoff two very useful generalisations, one relating to continuity of current in conductors, the other to the application of Ohm's law to complex arrangements of conductors. These generalisations are put into the form of two laws, known as Kirchhoff's laws, which are,

(i) The algebraic sum of the currents which meet at any point is zero.

(ii) In any closed circuit, the algebraic sum of the products of the current and resistance of each part of the circuit is equal to the electromotive force in the circuit.

By the application of these two laws, many problems on the currents in a network of conductors may be solved, and the resultant resistance of the network found.

From the first law we see that in the case of a number of conductors meeting at a point (Fig. 53 (i)) the relation

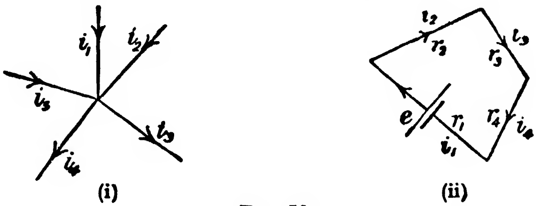


FIG. 53.

$$i_1+i_2-i_3-i_4+i_5+\ .\ .\ .\ =0$$

holds between the currents, the positive sign being given to those which flow towards the point and the negative sign to those which flow away from it. We shall see later that the first law is an expression of the fact that when the currents in a conductor are steady there is no accumulation of electricity anywhere.

The second law applied to a circuit such as Fig. 53 (ii) leads to the equation

$$i_1r_1+i_2r_2+i_3r_3+i_4r_4=e.$$

If the circuit is a mesh of a network, it may happen that one or more of the currents is negative, and this fact must appear by a suitable change of sign in the equation.

Wheatstone's Net.—The most important application of Kirchhoff's laws is in connection with the problem of the *Wheatstone's Net*, an arrangement of conductors used very widely for the practical comparison of resistances.

Four conductors, whose resistances are respectively r_1, r_2, r_3 and r_4, are arranged as in Fig. 54 (i). A current i enters at A and divides between the two branches ABC and ADC, the two parts uniting at C. If the potential difference between B and D

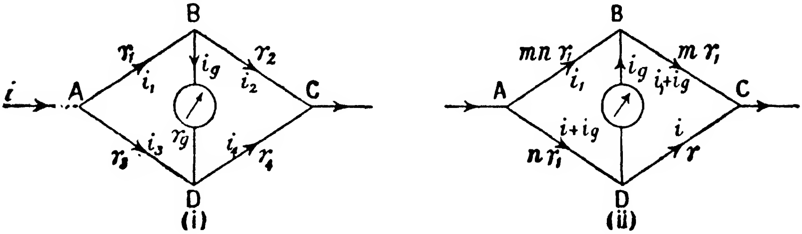


FIG. 54.

is zero, then, on joining B to D by a conductor, there will be no current in this conductor. This condition is rendered obvious by including a galvanometer in the circuit between B and D. Assuming currents i_g, i_1 and i_2 to flow in BD, AB and BC respectively, Kirchhoff's first law gives, for the point B,

$$i_1=i_2+i_g.$$

But with zero deflection of the galvanometer, $i_g=0$, and $i_1=i_2$. Similarly, $i_3=i_4$.

Applying Kirchhoff's second law to the circuit ABD, we have

$$i_1r_1+i_gr_g-i_3r_3=0,$$

and for the circuit ABCD,

$$i_1r_1+i_2r_2-i_4r_4-i_3r_3=0.$$

Remembering that $i_g=0$, and $i_1=i_2$, $i_3=i_4$, these two equations become—

$$i_1r_1=i_3r_3$$

$$i_1(r_1+r_2)=i_3(r_3+r_4)$$

$$\therefore \frac{r_1}{r_1+r_2}=\frac{r_3}{r_3+r_4}$$

$$\text{or,}\quad \frac{r_1}{r_2}=\frac{r_3}{r_4}.$$

Thus the condition for zero galvanometer deflection is that the four arms of the Wheatstone's bridge shall have resistances whose values form a proportion.

Sensitiveness of Bridge.—Although the bridge condition may

be fulfilled, it does not follow that this condition is readily detected. If, for example, r_3 is too large to satisfy the proportional condition, a current will flow through the galvanometer in the direction shown in Fig. 54 (i), and if too small the current will flow through the galvanometer in the reverse direction. When the departure of r_3 from its proper value is small, the current in the galvanometer will be small, and may be undetected unless the galvanometer is sensitive. Thus the tangent galvanometer is quite unsuitable for use with the Wheatstone's bridge, as a comparatively large current is necessary to produce an appreciable deflection. It follows that the more sensitive the galvanometer employed, the more accurately can the bridge balance be adjusted.

An examination of Fig. 54 (i) shows that the same proportion between the resistances is required to produce a balance, if the current entered at B and left at D, while the galvanometer is connected between A and C. In either case $r_1r_4=r_2r_3$, but it does not follow that the two possible arrangements are equally sensitive.

The question of the sensitiveness of the bridge has been treated by Prof. H. L. Callendar [1] in the following manner :—

The resistance to be measured is r, and is placed in the arm DC of the bridge (Fig. 54 (ii)), and the other arms have resistances mr_1, nr_1 and mnr_1, as shown. If i is the current in r and i_g the current in the galvanometer, and i_1 that in mnr_1, that in nr_1 is $i+i_g$, and that in mr_1 is i_1+i_g.

For the bridge condition to be satisfied, $r=r_1$, and it then follows, as on p. 66, that $i_g=0$; also $i=mi_1$. If, however, this condition is not fulfilled, the quantity $r-r_1$ is the most natural measure of the want of balance. The corresponding current in the galvanometer can then be calculated.

Applying Kirchhoff's second law to the circuits ABD and BDC in turn gives,

$$mnr_1i_1-r_gi_g-nr_1(i+i_g)=0$$
$$ri-mr_1(i_1+i_g)-r_gi_g=0,$$

which may be written—

$$mnr_1i_1-r_gi_g-nr_1i-nr_1i_g=0$$
$$-mr_1i_1-r_gi_g+ri-mr_1i_g=0.$$

Multiplying the latter by n and adding—

$$-(1+n)r_gi_g+ni(r-r_1)-(1+m)nr_1i_g=0,$$

or,
$$\frac{i_g}{i}=\frac{r-r_1}{\frac{1+n}{n}r_g+(1+m)r_1}.$$

[1] H. L. Callendar, *Proc. Phys. Soc.*, XXII, Part II, p. 220. 1910.

The quantity i_g/i is a useful expression for the sensitiveness of the bridge, as it gives the ratio of the current in the galvanometer to the main current produced by the battery, in terms of the want of balance of the bridge $(r-r_1)$ and the ratios m and n. It is independent of the resistance of the battery, and increases as the resistance of the galvanometer is made smaller, although the advantage of small galvanometer resistance may be lost if the sensitiveness of the galvanometer itself is unduly diminished.

For a given want of balance, $r-r_1$, the ratio i_g/i increases when n increases and when m diminishes, a limit being reached when $n=\infty$ and $m=0$, in which case,

$$\frac{i_g}{i}=\frac{r-r_1}{r_g+r_1}.$$

This condition is impracticable, because if $n=\infty$, the currents in the bridge would be zero, since the resistances in AB and AD are both infinite; and if $m=0$, the branch ADC and the galvanometer are short-circuited.

There is no great advantage in making n very large or m very small, for if $m=n=1$,

$$\frac{i_g}{i}=\frac{r-r_1}{2(r_g+r_1)},$$

and the sensitiveness is still half as great as in the ideal limiting case. It is of the greatest importance to note that n should never be very small, nor should m be very great, for in either case the sensitiveness is very much reduced.

The equation for the sensitiveness of the bridge may be written in the form

$$\frac{i_g}{r-r_1}=\frac{i}{\frac{1+n}{n}r_g+(1+m)r_1},$$

and it appears from this that the galvanometer current for a given want of balance of the bridge might be increased indefinitely by increasing i. The limit to i, the current in the resistance to be measured, is determined by the heating effect of this current and the consequent change of resistance due to it. The disturbance due to this effect is particularly important when the material of the resistance to be measured differs from that of the resistances forming the other three arms of the bridge. Prof. Callendar gives as a rule that, whenever there is a choice of arrangement, it is advisable to connect the battery so that the resistance in series with the resistance to be measured is greater than the resistance in parallel with it. This has the advantage of making the current in the resistance to be measured as small as possible, and at the same time making n greater than m.

For the most accurate determination of the proper balance of the Wheatstone's bridge, the greater the sensitiveness of the galvanometer the better, provided that in obtaining high sensitiveness of the galvanometer its resistance has not been unduly raised. The sensitiveness of the galvanometer depends upon the number of turns in the coil and the disposition of those turns (see p. 76). For a given current, the deflection is proportional to the number of turns in the coil, when the size and position of the turns is not altered by changing the number. Thus, starting with a single turn of wire whose cross-section is A, we see on replacing it by two turns of cross-section A/2 in series, and having the same radius, that we double the deflection for a given current. Or, if the single turn be replaced by s turns each of area A/s, all in series, the deflection for the same current is increased s times. But the resistance of the coil is increased s^2 times, that is, $r_g \propto s^2$. Therefore the deflection for a given current is proportional to $\sqrt{r_g}$. The deflection is also proportional to the current, and is hence proportional to the quantity $i_g\sqrt{r_g}$.

Referring now to the above equation for the sensitiveness of the bridge, and multiplying both sides by $\sqrt{r_g}$,

$$i_g\sqrt{r_g}=\frac{i(r-r_1)\sqrt{r_g}}{\frac{1+n}{n}r_g+(1+m)r_1}.$$

For given values of i and $(r-r_1)$, it follows that the galvanometer deflection is proportional to

$$\frac{\sqrt{r_g}}{\frac{1+n}{n}r_g+(1+m)r_1}.$$

In order to find the condition for this quantity to be a maximum, it may be differentiated with respect to r_g, and equated to zero, when we obtain,

$$(1+m)r_1=r_g\frac{1+n}{n}$$

$$r_g=\frac{n(1+m)r_1}{1+n}.$$

This gives the most efficient resistance for any given type of galvanometer for any arrangement of the Wheatstone's bridge. Unless n is made less than unity, or m greater than unity, the value of r_g lies between $r_1/2$ and $2r_1$, that is, the galvanometer should have the same order of resistance as r.

CHAPTER III

THE ELECTRIC CURRENT (*continued*)

Measurements

Galvanometers.—The name galvanometer is applied to those instruments which are used for the measurement or detection of very small currents. Although there is no absolute line of demarcation between these on the one hand, and ammeters which measure relatively large currents, and voltmeters which measure large differences of potential, on the other, still, the latter have as a rule a definite fixed scale, graduated to read amperes or volts, while the scale used with a galvanometer is not as a rule part of the instrument itself: and the value of the scale divisions, which for many purposes need not be known, is found for each experiment, when required, by some process of calibration.

In the case of the tangent galvanometer described in the last chapter, the expression for the current flowing in the coil is—

$$i=\frac{r\mathrm{H}}{2\pi n}\tan\theta,$$

and for reasons connected with the use of such a galvanometer, r is always great. This is no disadvantage so long as the current to be measured is large, but if the current is extremely small, r must obviously be as small as possible, and at the same time n must be as great as possible. Thus for high sensitiveness, many turns of small radius must be employed. Obviously, the number of turns cannot be indefinitely increased without making the radius of the turns so great that they are of very little value, and even act detrimentally by unduly increasing the electrical resistance of the galvanometer.

Having reached the limit of any increase of sensitiveness obtained by modifying the coil, we then turn our attention to the method of measuring the deflection. The scale and pointer method is only applicable for very rough instruments, and is replaced by the scale and mirror method described on p. 9. The scale being situated at some distance from the needle, and the rotation of the reflected beam of light being double that of the mirror, it follows that for deflections of a considerable number of

scale dimensions, the angle turned through by the needle is so small that tan θ may be replaced by θ itself, and this may be taken as proportional to the deflection in divisions on the linear scale. This is thoroughly justified if the reading is near the middle of the scale, but if the deflections are large the scale must be directly calibrated.

The calibration of a galvanometer scale may be most easily effected by means of a standard cell and a high resistance box. The box, the cell and the galvanometer are joined in simple series and the resistance adjusted until the deflection is the greatest allowable. The resistance and E.M.F. being known, the current is calculated (current $=\frac{e}{r}$). On increasing the resistance by suitable steps, other readings of the deflection may be obtained, and the corresponding current calculated, and the result may very suitably be represented by means of a curve, deflections being abscissæ and milli-amperes (thousandths of an ampere) or micro-amperes (millionths of an ampere) ordinates.

If the controlling magnetic field, H, which brings the needle into the plane of the coil, when no current is flowing, is the horizontal component of the earth's field, the range of sensitiveness of any given galvanometer is restricted; the only changes that can be made are effected by altering the coils or the method of reading the deflection; but if H can be varied, we have a great extension of the range of usefulness of the instrument. This is usually accomplished by means of a controlling magnet M (Fig. 55). Thus, if H be increased by lowering M, its S pole pointing north, and its N pole pointing south, we see that a given current will produce a less deflection, but if the N pole of M point north, or if instead M be raised, the controlling field H is weakened, and the given current will produce a greater deflection. Thus the sensitiveness may be considerably altered by means of the controlling magnet.

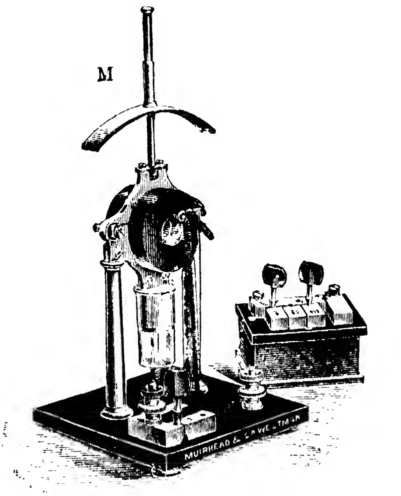


Fig. 55.

The use of a controlling magnet has the further advantage that the spot of light may by means of it be easily adjusted to the zero of the scale.

Another device for increasing the sensitiveness is to arrange the suspended magnetic system in such a way that the controlling field exerts an extremely small turning effect upon it.

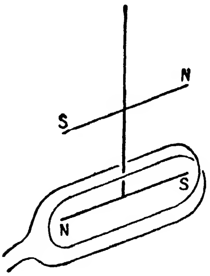


FIG. 56.

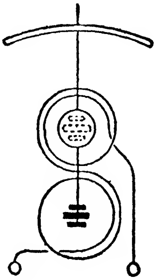


FIG. 57.

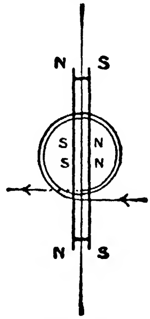


FIG. 58.

Two magnets are connected rigidly together in such a way that the couples exerted by the controlling field upon them are opposite and very nearly equal. They must not be exactly equal and opposite, in which case the system is said to be *astatic*, or the control would be zero, and the arrangement would be unworkable. The coil only surrounds one of the needles (Fig. 56), so that the deflecting couple acts on one of them only. In the Kelvin type of instrument two coils are employed, one surrounding each magnet, the coils being so connected that the couples due to the current both turn the systems of needles in the same direction (Fig. 57). A further arrangement due to Professor Broca, modified by Dr. Harker, is shown in Fig. 58. Two steel wire magnets are rigidly attached together vertically, one of which has N poles at its end, and consequent S poles in the middle, and the other S poles at the ends with consequent N poles in the middle. This arrangement allows powerful magnets to be used without having a large moment of inertia for the moving system.

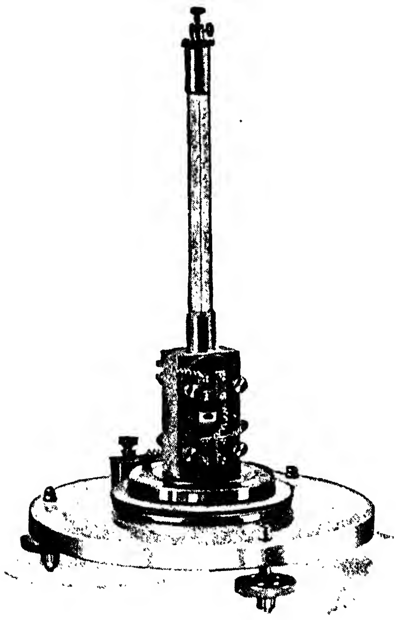


FIG. 59.

Fig. 59 illustrates a galvanometer of the Kelvin type.

The **Suspended Coil** galvanometer has largely superseded the galvanometer with the suspended magnet, for two important reasons—the instrument is not susceptible to disturbance by varying external magnetic fields, and the suspension is much stronger. On the contrary, its sensitiveness cannot be varied at will, and the damping is usually much greater. This latter is not always a disadvantage ; in fact, for deflection measurements and those involving the finding of an electric balance as in the case of the Wheatstone's bridge, it may be a great advantage, but in measurements of the ballistic type, as we shall see in Chapter VIII, it is a serious objection.

The principle involved in the use of this type of galvanometer must be deferred to Chapter VIII, but on general grounds it may be seen that if a magnet experiences a couple due to the current in a coil, the coil experiences an equal and opposite couple, so that if the magnet be fixed and the coil suspended, the latter will be deflected when a current flows in it, and further, the couple is proportional to the current. To make this change, a radical alteration in design is necessary. Since the magnet is fixed we may make it as large and heavy as we please, while on the other hand, the coil, being suspended, must be as light as possible.

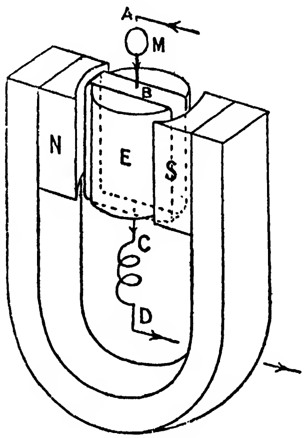


FIG. 60.

A typical arrangement for this form of galvanometer is shown in Figs. 60 and 61. A permanent magnet, usually of the horse-shoe type, has soft iron pole-pieces N and S. The coil B, which is commonly rectangular, is suspended between these pole-pieces by a fine phosphor-bronze strip, which also serves as a conductor to bring in the current. After passing round all the turns of the coil, the current passes out by means of the second phosphor-bronze strip CD. When the current flows, the coil experiences a couple, and will rotate until the couple due to the twist in the suspension becomes equal and opposite to the deflecting couple. The couple due to this twist in the suspension is proportional to the angle of twist itself, and the coil should be always in a magnetic field of the same strength. To ensure this, the soft iron cylinder E is situated between the pole faces, and serves the double

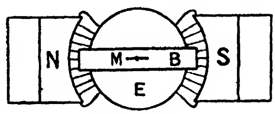


FIG. 61.

purpose of making (by the presence of the poles upon it) the field stronger and also of making it radial, as shown in the plan (Fig. 61). We shall see in Chapter VIII that the couple acting on the coil is proportional to the current i, to the number of turns n, and to the area of each turn A, and to H, the strength of field in which the sides of the coil are situated.

$$\therefore \text{Couple} \propto i\mathrm{A}n\mathrm{H}.$$

The controlling couple is proportional to the deflection θ, and to the couple c for unit twist (one radian) of the suspension, the upper end being fixed. Thus, for equilibrium

$$i\mathrm{A}n\mathrm{H} \propto c\theta, \text{ or, } i \propto \frac{c}{\mathrm{A}n\mathrm{H}}\theta.$$

Hence for a well-designed instrument the current is proportional to the deflection itself, since the field is certainly of con-

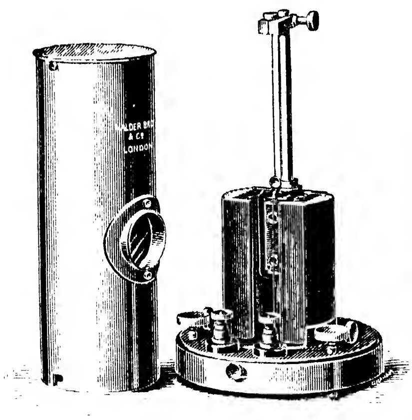


FIG. 62.

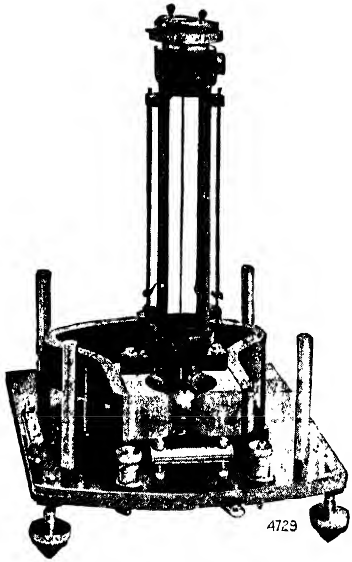


FIG. 63.

stant value until θ attains much larger value than that occurring in any probable experiment. To increase the sensitiveness, c may be diminished, and A, n and H increased. c cannot be diminished without limit, as the suspension must be strong enough to carry the coil. It is usually made of phosphor-bronze, as this has a tensile strength approaching that of steel and is not readily oxidisable. The coil is made as light as possible by constructing it of thin high-conductivity insulated copper wire, but it may be noticed that we can never use as great a number of

turns of wire in the coil as in the case of the suspended magnet instrument, and for this reason the suspended coil galvanometer has generally a lower resistance. The loss of sensitiveness due to having fewer turns is, however, made up for by the very high value of the magnetic field employed. The permanent magnets are generally built up of several hard-steel horseshoe magnets to ensure permanence and great strength of field.

Galvanometers of the suspended coil type are frequently called d'Arsonval galvanometers after their inventor.

Einthoven's String Galvanometer.—A silvered quartz fibre AB is stretched vertically between the poles NS of an electromagnet (Fig. 64). On observing the fibre through the channel CD by means of a microscope, a displaced image of the fibre is seen when a current flows in it (Chap. VIII). The displacement on the eyepiece scale is a measure of the current. The instrument may also be used as an oscillograph (Chap. X) by allowing the fibre to move in front of an illuminated horizontal slit and projecting the image on a moving photographic film. The arrangement is very sensitive, currents of 10^{-12} ampere being measurable. **Also its frequency of vibration may be high, so that** its movement is a faithful representation of the variation of current in it.

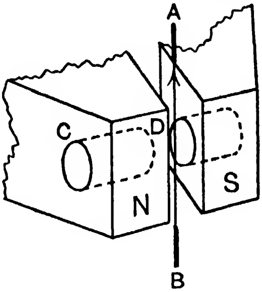


FIG. 64.

In the **Thermo Galvanometer** due to Duddell, the heating produced in a fine wire is measured by an arrangement similar to that used in the Boys' radio-micrometer. A loop of silver wire hangs between the poles of a permanent magnet and the loop ends in two little pieces, one of antimony and the other of bismuth. These are in contact at their lower extremities, and are situated over the heater H (Fig. 65), which carries the current to be measured. When a temperature difference exists between the lower Bi-Sb junction and the rest of the circuit, a current proportional to this temperature difference flows in the loop, which then rotates until the torsion in the suspending quartz fibre F brings it to rest. The rate of production of heat in H is proportional to the square of the current, and this is also found to be proportional to the deflection. The heaters are made of various resistances from 4 ohms to 1000 ohms, those of lower resistance being metal wires and those of high resistance consisting of a deposit of platinum on quartz.

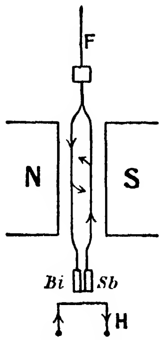


FIG. 65.

With a scale at a distance of one metre, a current of 110 micro-amperes gives a deflection of 250 millimetres, using a 1000-ohm heater, and with a 4-ohm heater a p.d. of 7 millivolts gives a deflection of 250 millimetres. The great advantage of the instrument lies in the fact that it is equally applicable to the measurement of direct or alternating currents, and on being calibrated for one, may be used to measure the other. It is therefore useful for the measurement of high-frequency oscillating currents of a few micro-amperes.

Sensitiveness of Galvanometer.—The *Figure of Merit* of a galvanometer is the current which will produce a deflection of one scale division. This depends upon the distance of the scale from the galvanometer and the size of division, and thus for facility in comparing different galvanometers it is usual to employ a scale of millimetres, placed at a distance of one metre from the mirror of the galvanometer. Provided that the deflection is not too large, we see that the current is proportional to the deflection, both for the suspended magnet and the suspended coil type of galvanometer, and therefore we can obtain the figure of merit by observing the deflection for a given current, as in the calibration described on p. 71. Dividing the current by the deflection the figure of merit is obtained.

There is considerable difficulty in making a comparison of the efficiency of different galvanometers, since this depends so much upon the use to which the instrument is to be put. It might happen that an extremely small current produces a considerable deflection, and yet the resistance of the galvanometer may be so great that it is unsuitable for some purposes, as, for example, making measurements by the Wheatstone's bridge. If a constant p.d. be maintained between the terminals of a given galvanometer, the deflection is independent of the resistance of the coils, provided that all the turns are equally effective in producing magnetic field, for the field is then proportional to the number of turns. But, on the other hand, the resistance is also proportional to the number of turns, so that the current corresponding to the given p.d. will vary inversely as the number of turns, and the deflection, being proportional to the field and the current, will remain unaltered.

Again, the moment of inertia of the suspended part does not affect the steady deflection for any given current, but it does affect the period of swing and the time for which the needle goes on swinging, and a large moment of inertia may, therefore, render a sensitive galvanometer unsuitable for bridge work and for measurements in which a steady deflection is to be obtained rapidly. And, further, the suspended magnet galvanometer is usually provided with a moveable controlling magnet, so that

the controlling field, and therefore the sensitiveness, may be varied between wide limits.

For these reasons it has been suggested that the sensibility of a galvanometer should be defined as the number of scale divisions deflection for a current of one micro-ampere when the scale is at a distance of 1000 scale divisions from the mirror, reduced to the corresponding value for the same rate of expenditure of energy when the resistance of the galvanometer is one ohm and the period of vibration one second.

Let the deflection be θ scale divisions for a current of one micro-ampere when the resistance is R ohms and the time of vibration T seconds. The rate of working is now $R\times10^{-12}$ watts (since current is one micro-ampere$=10^{-6}$ amperes), and therefore to maintain the same rate of working with the resistance changed to one ohm, the current would be $\sqrt{R}$ micro-amperes, and the deflection for one micro-ampere would, under the new conditions, be $\frac{\theta}{\sqrt{R}}$. To reduce the deflection to correspond to a period of vibration of one second, remember that $T=2\pi\sqrt{\frac{I}{MH}}$ in the case of a vibrating magnet, where H is the controlling field;

$$\therefore\ H\propto\frac{1}{T^2}.$$

But deflection $\propto\frac{1}{H}$ (p. 71);

$$\therefore\ \text{Deflection}\propto T^2.$$

Hence to find the deflection if the periodic time were reduced to one second, we must divide by T^2.

$$\therefore\ \text{Sensibility}=\frac{\theta}{T^2\sqrt{R}}.$$

A similar process of reasoning applies to the suspended coil galvanometer, for in this case $T=2\pi\sqrt{\frac{I}{c}}$, where c is the restoring couple for unit twist in the suspension, and deflection $\propto\frac{1}{c}$.

$$\therefore\ \text{Deflection}\propto T^2.$$

Galvanometer Resistance.—The resistance of the galvanometer may vary between wide limits, but that of the suspended magnet type is, in general, greater than that of the suspended coil. The latter may have any resistance up to 2000 or 3000 ohms, but in the case of the suspended magnet instrument it reaches in

exceptional cases as much as 300,000 ohms, although 4000 to 6000 ohms is a usual value.

There are many methods of measuring the resistance of a galvanometer, but undoubtedly the most accurate is to clamp the coil and treat the galvanometer as an ordinary conductor, and measure its resistance by comparison with a standard, using the metre bridge or Post Office box. This involves the employment of a second galvanometer, which presents no difficulty in a physical laboratory; but there are several methods of determining the resistance without the use of this second galvanometer. For example, in Kelvin's method, described on p. 94, the galvanometer itself occupies one arm of a Wheatstone's bridge. The following two simple methods will give the resistance approximately.

If the galvanometer has a low resistance G, connect it in series with a cell of E.M.F. E volts, and a resistance box in which resistance R_1 ohms is used.

Then, $I_1 = \frac{E}{G+R_1}$, assuming that the resistance of the cell is negligible. Now change the resistance in the box to R_2 ohms, when

$$I_2 = \frac{E}{G+R_2}.$$

$$\therefore \frac{I_1}{I_2} = \frac{G+R_2}{G+R_1}, \quad G = \frac{R_2 I_2 - R_1 I_1}{I_1 - I_2}$$

I_1 and I_2 are proportional to the deflections in the case of a reflecting galvanometer, and to the tangents of the deflections in the case of a tangent galvanometer. The method is applicable when the galvanometer is not very sensitive, but in the case of a delicate instrument, R would have to be so great in order to produce a reasonably small deflection, that the method is useless, and the next method may be employed. A conductor of resistance r_1 is placed in parallel with the galvanometer and another resistance R_1 in series (Fig. 66). Then R_1 being usually very great in comparison with the resistance of the cell,

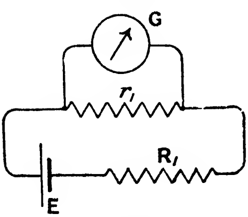


FIG. 66.

$$\text{Current in galvanometer, } I = \frac{E}{R_1 + \frac{r_1 G}{r_1 + G}} \cdot \frac{r_1}{r_1 + G}.$$

If now r_1 be changed to r_2, and R_1 be changed to the value R_2,

such that the current I in the galvanometer is the same as before,

$$\frac{E}{R_1+\frac{r_1G}{r_1+G}}\cdot\frac{r_1}{r_1+G}=\frac{E}{R_2+\frac{r_2G}{r_2+G}}\cdot\frac{r_2}{r_2+G}$$

$$\frac{r_1}{R_1r_1+R_1G+r_1G}=\frac{r_2}{R_2r_2+R_2G+r_2G}$$

$$G=\frac{r_1r_2(R_1-R_2)}{r_1R_2-r_2R_1}$$

Galvanometer Shunts.—Resistances are sometimes placed in parallel with the galvanometer to reduce the sensitiveness, by offering an alternative path to the current, so that only a fraction of it goes through the galvanometer. Thus, if the galvanometer resistance be G, and that of the shunt S (Fig. 67), then for the

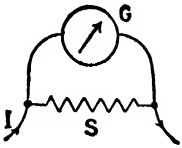


FIG. 67.

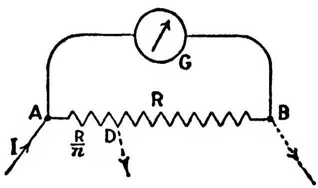


FIG. 68.

current I flowing in the main circuit, that in the galvanometer is $I\cdot\frac{S}{G+S}$. With the older galvanometers, boxes of shunts were supplied, whose resistances were respectively $\frac{1}{9}$, $\frac{1}{99}$, and $\frac{1}{999}$ of that of the galvanometer, and in this case $\frac{S}{G+S}$ has the value $\frac{1}{10}$, $\frac{1}{100}$ or $\frac{1}{1000}$, so that the sensitiveness of the galvanometer is reduced 10, 100 or 1000 times by the use of the proper shunt. Each galvanometer had its own shunt, which sometimes led to great inconvenience, and hence the advantage of the Ayrton and Mather Universal Shunt, which may be applied to any galvanometer.

The high resistance R is placed in parallel with the galvanometer, and the current I enters at A and leaves at B, Fig. 68,

$$\text{current in galvanometer}=I\frac{R}{G+R}.$$

If now the point at which the current leaves be transferred to

D, such that resistance AD$=\frac{1}{n}$ (resistance AB), the two circuits AD, $\left(\frac{R}{n}\right)$, and AGBD, $\left(G+R-\frac{R}{n}\right)$, are in parallel, and

$$\text{current in galvanometer}=I\frac{\frac{R}{n}}{\frac{R}{n}+G+R-\frac{R}{n}}=\frac{I}{n}\,\frac{R}{G+R},$$

that is, it is $\frac{1}{n}$ of the previous current.

By having a number of points such as D, for which $n=1$, 10 100, 1000, etc., the effect of the shunt may be conveniently varied. Fig. 69 shows the arrangement sometimes adopted in the Ayrton and Mather shunt: G and G are connected to the galvanometer and L L are the circuit terminals. With the rotating arm as shown, the value of the shunt is $\frac{1}{10}$.

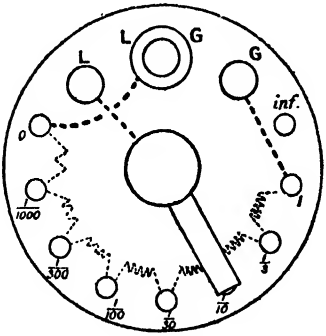


FIG. 69.

It should be noted that in moving the point of contact from B to D (Fig. 68), the effective resistance between the leads falls from

$$\frac{GR}{G+R} \quad \text{to} \quad \frac{\frac{R}{n}\left(G+R-\frac{R}{n}\right)}{\frac{R}{n}+G+R-\frac{R}{n}}$$

$$i.e.\ \text{to}\ \frac{R}{n}\cdot\frac{\left(G+R-\frac{R}{n}\right)}{G+R},$$

and the main current may be thereby altered. For the main current to be unchanged, these values of the resistance must be the same,

$$\therefore\ G=\frac{G+R-\frac{R}{n}}{n},\ \text{or},\ R=nG.$$

Thus for any given value of the shunt, say $\frac{1}{10}$, the effective resistance of the circuit is unchanged on employing the shunt, provided that the resistance AB is ten times the galvanometer resistance. The sensitiveness of the galvanometer is reduced on attaching the universal shunt, in the ratio $1:\frac{R}{G+R}$, that is by

$100\left(1-\frac{R}{G+R}\right)$ or $100\frac{G}{G+R}$ per cent.; but this is $100\frac{G}{G+nG}$, when $R=nG$, or $\frac{100}{1+n}$. In the case when $n=10$, this amounts to 9·1 per cent., and when $n=100$ or 1000 it is unimportant. The resistance of the whole shunt should therefore be at least 100 times that of the galvanometer.

Voltmeters.—Although the variety of voltmeters in use is very great, and many different principles are used in their construction, we may broadly distinguish between the electromagnetic type, which is simply a galvanometer of low sensibility having a fixed scale graduated to read volts, and the electrostatic type, which is a modified form of electrometer, the description of which will be left to Chapter V. But one characteristic is common to them all; they must have very high resistance, the reason being that

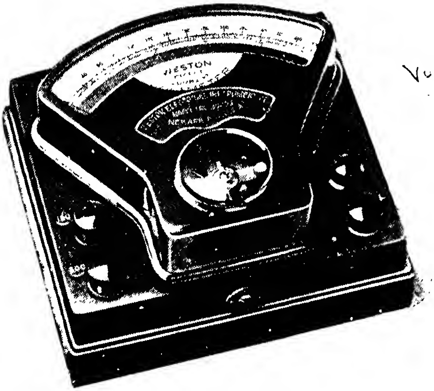


FIG. 70.

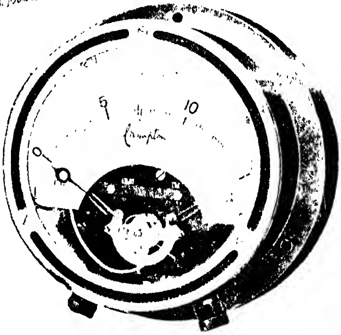


FIG. 71.

in connecting them between the two points whose p.d. is required, they must take only an infinitesimal current, so that no appreciable disturbance of the circuit is produced, and the heat produced by the current in the instrument itself shall be negligible. In the electromagnetic type of instrument the coil is situated in a strong radial magnetic field, as in the case of the galvanometer, but the control is produced by a spiral spring usually made of phosphor-bronze, and the coil turns upon two jewelled pivots (Fig. 70). A suitable high resistance is connected permanently in series with the galvanometer movement.

In Fig. 71 is shown another instrument of the moving coil type, which is supplied either as a millivoltmeter or as a milliamperemeter according to the scale attached. It also serves the purpose of a galvanometer when great sensitiveness is not

required, the scale, in this case, being of the lowest range. The resistance of the moving part is 5 ohms and the reading is one large division for 1 milliampere or, as a voltmeter, one large division for 5 millivolts.

Ammeters.—The same galvanometer movement that is used for the voltmeter may also be used for an ammeter. In this case, instead of the high series resistance, a low resistance shunt is employed. The shunt AB (Fig. 72) is placed in the circuit in which the current to be measured is flowing and the galvanometer movement G joined to A and B. For example, if the galvanometer has a deflection of one scale division for 1 milliampere and its resistance is 50 ohms, then if R is the resistance of AB, and a current of 1 ampere flows in the main circuit, $\frac{R}{50+R}$ is the current in G. When this is one milliampere

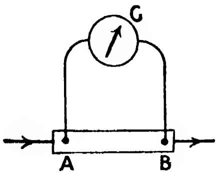


FIG. 72.

$$\frac{R}{50+R}=0{\cdot}001, \quad 0{\cdot}999R=0{\cdot}05,$$

$R=0{\cdot}05005$ ohm. With this resistance for the shunt, each scale division corresponds to a main current of one ampere.

Many instrument makers place the shunt in the case of the instrument, and others supply sets of shunts, which enormously increase the range. Fig. 73 shows a thousand-ampere shunt for switchboard mounting (Messrs. Everett, Edgcumbe & Co., Ltd.).

In all cases the resistance of an amperemeter must be very small; in the first place so that no disturbance is made in the circuit in which it is placed, and secondly because the heat produced in the instrument must, in spite of the large current, be negligible.

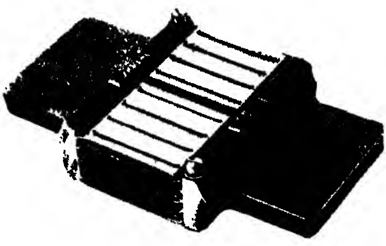


FIG. 73.
[*By courtesy of Everett Edgcumbe & Co., Ltd.*]

Hot Wire Instruments.—The heat produced by the passage of the current has also been employed in the operation of ammeters and voltmeters. The current passes through a fine wire which is thereby heated. The expansion is observed by some mechanical multiplying arrangement. The Cardew voltmeter is an example of this form of instrument. The advantage lies in the fact that the direction of movement of the pointer is independent of the direction of the current, and the instrument

can therefore be used for measuring alternating currents. The hot wire instruments, however, are very liable to change of zero, and the fact that the heating is proportional to the square of the current renders the scale uneven, and further, the multiplying mechanism is a frequent source of uncertainty.

Soft Iron Instruments.—*Soft Iron* or *moving iron* ammeters also are constructed, which read satisfactorily on both continuous and alternating current circuits. The principle of such instruments is illustrated in Fig. 74. The current to be measured passes round a coil wound upon a brass cylinder, and therefore gives rise to a magnetic field parallel to the axis of the coil. Two soft iron bars, A and B, are situated in this field, with their axes parallel to it, and are therefore magnetised when the current flows in the coil, to a degree depending upon the strength of the current. A is fixed to the brass cylinder and B is part of a light framework pivoted upon centres O situated in the axis of the coil. When A and B are magnetised, the ends situated together are poles of the same kind, that is both N or both S, and the two bars therefore repel each other. The force of this repulsion is proportional to the product of the two pole strengths, and each pole depends upon the strength of current and hence the repulsion depends upon the square of the current. The control may be gravitational, the moving system being so balanced that the pointer S is at zero on the scale with no current, or by a spiral spring. The relation between deflection and current is complicated, so that the scale, which is usually more open in the middle than at its ends, must be calibrated by comparison with an ampere balance or other standard instrument.

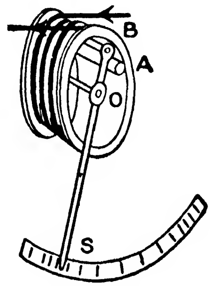


FIG. 74.

Since the deflection depends upon the square of the current, its direction is independent of that of the current, and an alternating current will therefore produce a deflection. This type of ammeter, calibrated to read alternating currents, is very widely used.

Resistances.—For purposes of electrical testing, it is usual to arrange the standard sets of resistances in boxes, and the variety of ways in which this is done is very great. One very common arrangement is shown in Fig. 75. The wire whose resistance is approximately that required, is soldered one end to each of the brass blocks, which are themselves screwed

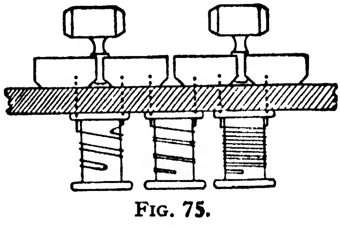


FIG. 75.

to the ebonite base. The wire, which is usually of manganin on account of its small temperature coefficient, is then doubled and wrapped round the bobbin as shown, and afterwards soaked in melted paraffin wax. The plugs are ground to fit into conical holes between the brass blocks, so that on inserting any given plug the corresponding resistance coil is short circuited.

Various types of resistance box are illustrated in Figs. 76 to 78.

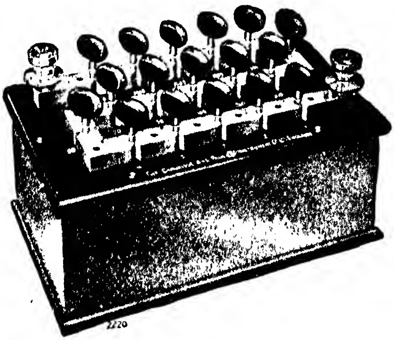


FIG. 76.

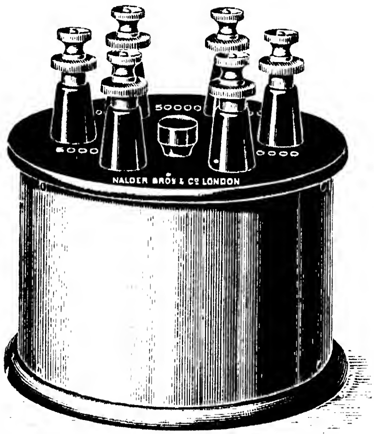


FIG. 77.

Owing to the change of resistance with temperature, resistance coils used for accurate purposes are provided with a thermometer to enable the temperature to be observed at the time of experi-

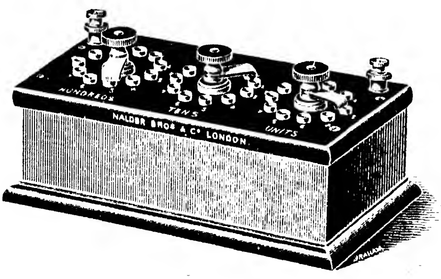


FIG. 78.

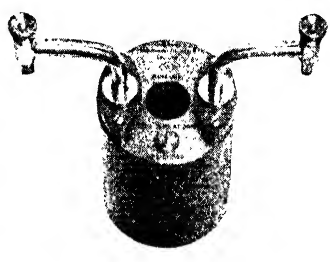


FIG. 79.

[*By courtesy of the Cambridge Instrument Co., Ltd.*]

ment. In Fig. 79 is shown a standard resistance coil whose terminals are stout copper conductors, which are amalgamated at the tips to ensure good electrical contact with the mercury cups in which they rest. The whole is immersed in a liquid bath to

maintain steadiness of temperature, and the hole enables a thermometer to be inserted so that the bulb shall be as near as possible to the resistance wire itself.

Measurement of Resistance.—The easiest method of comparing resistances is that of *simple substitution*. A cell of steady E.M.F., which need not be known, is placed in series with a galvanometer, a resistance box, and the resistance to be determined. The current is adjusted to a suitable amount either by changing the resistance in the box or by shunting the galvanometer. On removing the unknown resistance from the circuit, the current is increased, and may be brought back to its original value by increasing the resistance in the box by an amount equal to that of the resistance removed, which is therefore known.

The *deflection method* may be employed when the resistance to be measured is sufficiently high; the deflection θ being observed

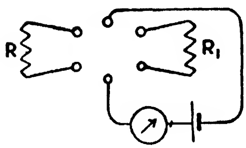


FIG. 80.

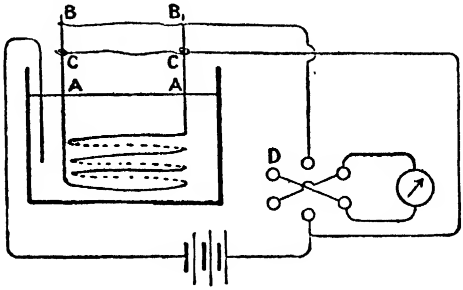


FIG. 81.

when the unknown resistance R is in circuit. This is then replaced by a known standard resistance R_1, and the deflection θ_1 observed (Fig. 80). If E is the E.M.F. of the cell, B its internal resistance, and G the resistance of the galvanometer—

$$I=\frac{E}{R+B+G}, \quad I_1=\frac{E}{R_1+B+G},$$

$$\frac{I}{I_1}=\frac{R_1+B+G}{R+B+G}=\frac{\theta}{\theta_1}.$$

If R and R_1 are very great in comparison with B and G, $\frac{R}{R_1}=\frac{\theta_1}{\theta}$.

The method is only employed in measuring very great resistances, such as that of the insulation of an electric cable, which is generally of the order of millions of ohms. The cable is immersed in a tank of water, and the current passed from the core B to the water of the tank, passing through the insulating layers of the cable. It will be seen that a leakage may occur over the surface of the insulation from A to B (Fig. 81), this current causing the

results obtained to be false. To get over this difficulty Price[1] has suggested that a wire, C, be bound round the insulation near the end of the cable and so connected to the rest of the circuit that this disturbing current will not pass through the galvanometer. The current in the galvanometer may conveniently be reversed by means of the commutator, shown at D, and the mean of the readings on each side of the zero, taken as the deflection.

A more direct method of measuring resistance is sometimes employed when the conductor is carrying current, as, for example, in the case of an incandescent lamp, the resistance of which, when hot, differs greatly from that when cold. The ammeter A (Fig. 82) indicates the current in the lamp, and the voltmeter V the potential difference between its terminals. By division we can obtain the resistance in ohms. The method does not admit of very great accuracy, but this is rarely required in such a case.

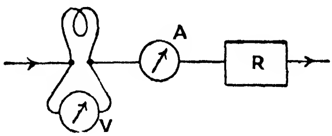


FIG. 82.

On altering the current by means of the variable resistance R, the resistance of the lamp for different currents may be found, and if, instead of dividing the p.d. by the current, we multiply them together, the result obtained is the power in watts absorbed by the lamp. The candle-power for different currents may be found by the ordinary photometric methods. A carbon filament lamp has about 50 per cent. less resistance when hot than when cold, while for a metallic filament the resistance when hot is many times that when cold.

Wheatstone's Bridge.—The method of the *Wheatstone's net*, described in the last chapter, affords the most accurate and frequently used method of comparing resistances, and it has the great advantage that it is a null method, the galvanometer being required to detect a want of balance and not to measure a deflection.

Fig. 83 (i) is a diagrammatic representation of the Wheatstone's net, and Fig. 83 (ii) shows the arrangement used in the *Post Office box*. We have seen on p. 66 that when $\frac{P}{Q}=\frac{R}{S}$, the current in the galvanometer is zero. Hence, if the ratio $\frac{P}{Q}$ and the value of R in ohms be known, the resistance S can be calculated. In the Post Office box the arms P and Q usually each consist of three resistances of 10, 100 and 1000 ohms respectively, so that any

[1] W. A. Price, *Electrical Review*, vol. 37, p. 702. 1895.

with respect to the bridge by means of the thick copper conductors CDEF, which, for convenience, are mounted on one

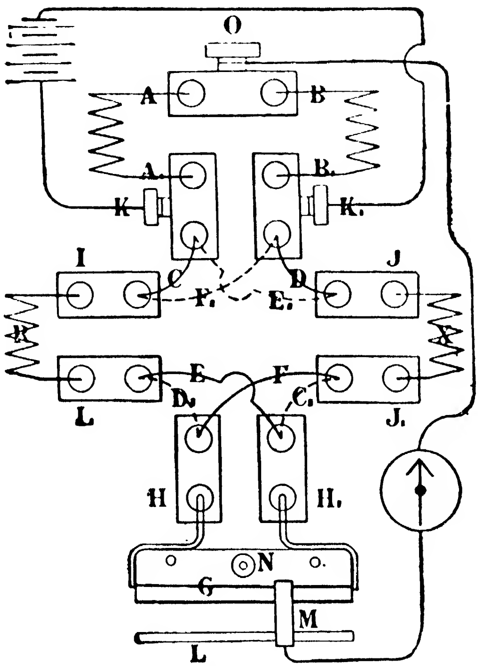


FIG. 87.
(From Henderson's "Practical Electricity and Magnetism.")

ebonite block, the rotation of which through 180° executes the required interchange.

Callendar and Griffiths Bridge.—This is another adaptation of the Wheatstone's bridge to a special purpose; in this case the measurement of temperature by the change in resistance of a platinum wire.

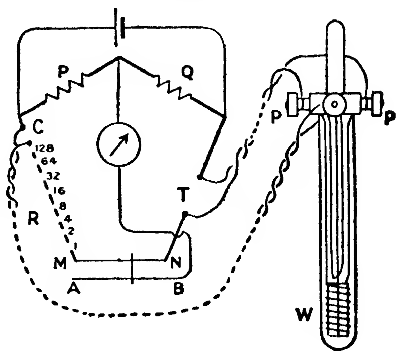


FIG. 88.

P and Q (Fig. 88) are the ratio arms which are adjusted to equality, and the platinum wire W is connected by internal leads to PP, and thence to the gap T of the bridge. An equivalent pair of leads joined together within the tube containing W, go to the gap C, and thus the disturbing effect of the leads on account of

their variation in temperature is eliminated. The two pairs are in practice kept close together, and their two resistances being equal are introduced in opposite arms of the bridge, which is adjusted for equality. The bridge wire MN is 50 cm. long, and at 1, 2, 4, 8, 16, 32, 64 and 128 are placed resistances whose values are 1, 2, 4 etc., times the resistance of 20 centimetres of MN. Since this has a resistance of about $\frac{1}{200}$ ohm per centimetre, the resistances of 1, 2, 4 etc., are therefore 0·1, 0·2 etc., up to 12·8 ohms, and with all of them in, W may have a total resistance of 25·5 ohms. The balance is obtained by moving the cross-piece which connects the wires AB and MN, all of which are made of the same material in order to avoid thermoelectric disturbances.

If the balance is obtained with the slider at a distance l from the middle of MN, and ρ is the resistance of a centimetre of MN, then, since the ratio arms are equal—

$$r+\mathrm{R}+m+\rho l=r+\mathrm{T}+m-\rho l,$$

where r is the resistance of the leads on either side, and m is the resistance of half of MN.

$$\therefore\ \mathrm{T}=\mathrm{R}+2\rho l.$$

The change of resistance of a platinum wire between 0° C. and 100° C. when the resistance at 0° C. is 12·8 ohms, is 5 ohms, which is equivalent to 0·05 ohm for 1° change. If, then, $\rho=0{\cdot}005$, l must change by $\dfrac{0{\cdot}05}{2\times0{\cdot}005}=5$ cm. to maintain a balance as the temperature rises 1°. Hence a millimetre of wire corresponds to a change of temperature of $\frac{1}{50}$ degree.

It was shown by Prof. H. L. Callendar [1] that the resistance of a pure platinum wire may be accurately represented by the relation

$$\mathrm{R}_t=\mathrm{R}_0(1+\alpha t+\beta t^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

over a very great range of temperature, where R_0, α and β are constants which may be found by measuring R_t at three different temperatures. For limited ranges, the term βt^2 may be neglected and the values of the resistance at 0° C. and at 100° C. found by the usual method of finding the fixed points of a thermometer; but for extended ranges the temperatures 0° C., 100° C. and the boiling point of sulphur at 760 mm. of mercury pressure (444·60° C.) are employed.

Prof. Callendar has also shown that if the temperature t_p on the platinum thermometer scale be calculated from the relation

$$t_p=100\frac{\mathrm{R}_t-\mathrm{R}_0}{\mathrm{R}_{100}-\mathrm{R}_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

[1] H. L. Callendar, *Phil. Trans.*, 178, p. 1. 1887.

where R_t, R_0 and R_{100} are the resistances of the thermometer at $t°$, 0° and 100° C. respectively, the difference $(t-t_p)$ between t_p and the temperature t on the gas thermometer scale is given by the relation

$$t-t_p=\delta\left(\frac{t}{100}-1\right)\frac{t}{100}$$

For, from (i), $R_{100}=R_0(1+100\alpha+10{,}000\beta)$,

or
$$\frac{R_{100}-R_0}{100R_0}=\alpha+100\beta$$

Again, from (ii), $R_t=R_0\left(1+\frac{R_{100}-R_0}{100R_0}t_p\right)$. Comparing this with (i)

$$\frac{R_{100}-R_0}{100R_0}t_p=\alpha t+\beta t^2$$

$$\therefore\ (\alpha+100\beta)t_p=\alpha t+\beta t^2$$

$$\frac{t_p}{t}=\frac{\alpha+\beta t}{\alpha+100\beta}$$

$$\frac{t_p-t}{t}=\frac{\beta(t-100)}{\alpha+100\beta}$$

$$t-t_p=\frac{-10^4\beta}{\alpha+100\beta}\left(\frac{t}{100}-1\right)\frac{t}{100}$$

and
$$\delta=\frac{-10^4\beta}{\alpha+100\beta}.$$

Griffiths has shown that for pure platinum δ has the value of very nearly 1·5. For any particular thermometer, t_p may be determined at the boiling point of sulphur, so that $t-t_p$ and therefore δ may be found for this thermometer.

Then if a curve connecting t_p and t be plotted, the correction to be added to t_p to obtain t for any value of t_p may be read upon it.

In order that the resistance between M and N shall be exactly $\frac{1}{4}$ ohm the actual resistance of the wire is slightly greater than this, and it is shunted with a fine wire whose length is adjusted until that of the combined resistance is exactly $\frac{1}{4}$ ohm. This does not in any way alter the point of balance, and ρ is now $\frac{1}{50}$ of combined resistance between M and N.

For temperatures up to 300° C. the thermometer consists of a platinum wire wound upon a mica frame and enclosed in a glass tube, but for higher temperatures the tube must be of glazed porcelain.[1]

[1] For experimental details of finding the fixed points and for calibrating the bridge, the student may with advantage consult "A Text Book of Practical Physics," by W. Watson.

Galvanometer Resistance (Kelvin).—The determination of the resistance of a galvanometer by means of the Wheatstone's bridge was first performed by Lord Kelvin, and is generally known as the Kelvin method. We have seen on p. 66 that the current in the galvanometer circuit (Fig. 83 (i)) is zero when $\frac{P}{Q}=\frac{R}{S}$, and consequently the currents in the remaining arms are independent of r_g. The galvanometer is therefore placed in the arm S, Fig. 89, when, on making the battery circuit, a deflection will be produced. This may be reduced to a readable amount by one of two methods: a resistance, T, may be introduced into the battery circuit, which reduces the whole current in the bridge, and therefore the sensitiveness of the test, or the spot of light may be brought back on to the scale by means of the controlling magnet in the case of a suspended needle instrument, or by twisting the torsion head in the case of a suspended coil instrument. In the case of a delicate galvanometer it may be desirable to combine the two processes. The resistances in the arms are then adjusted until the spot of light is brought to rest, and will still remain at rest whether the key K be open or closed. When this condition is attained $\frac{P}{Q}=\frac{R}{S}$, and S is, in this case, the resistance of the galvanometer.

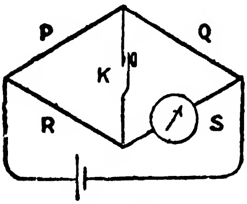


FIG. 89.

Battery Resistance (Mance).—A somewhat similar method due to *Mance* has been used for determining the resistance of a battery. The battery is placed in S, Fig. 90, and the key at K. When the key is closed, a current flows in the circuit, and we may imagine this current reduced to zero by an appropriate E.M.F. in K. This additional E.M.F. would not produce any current in g when $\frac{P}{Q}=\frac{R}{S}$, but the current in K being zero, it is immaterial whether the key be open or closed. Therefore the condition for the resistances in the arms to be proportional is that the current in the galvanometer is unaltered by opening or closing K. The method is not a good one, owing to the fact that the current in the galvanometer at the time of the test is large, and also that an unknown current is flowing in the battery. The resistance of the battery generally depends upon the current flowing, and this should be known, as in the method on p. 101.

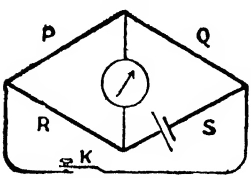


FIG. 90.

Low Resistances.—The Wheatstone's bridge is unsuitable for the comparison of very low resistances, for two reasons. The first has been discussed on p. 68, and refers to the want of sensitiveness of the bridge. The other important reason is that with low resistances, the connecting wires and the contacts at the terminals have resistances which are no longer negligible, and they may even be as great as, or greater than, the resistances to be compared, unless special precautions on this account are taken.

By the method of *direct deflection*, the two low resistances may be compared. A steady current is passed through the two conductors in series, and the galvanometer deflection produced by the p.d. between two points separated by a distance l_1 upon the first conductor (Fig. 91) compared with the corresponding deflec-

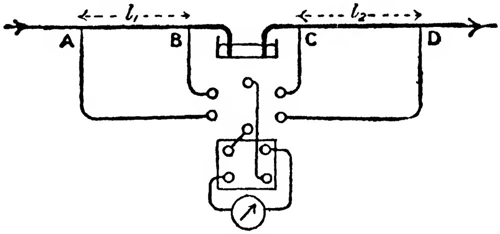


FIG. 91.

tion for the distance l_2 upon the second conductor. Then, if the galvanometer have a resistance which is high compared with that of the conductors under comparison (it is usually thousands of times as great), the current in the galvanometer will be inappreciable in comparison with that in the conductors. The current in the first case is proportional to E_1, the fall of potential over l_1, and the second to E_2, the p.d. over l_2. Then if R_1 and R_2 are the corresponding resistances, and θ_1 and θ_2 the deflections—

$$\text{Current}=\frac{E_1}{R_1}=\frac{E_2}{R_2}, \qquad \therefore \frac{R_1}{R_2}=\frac{E_1}{E_2}=\frac{\theta_1}{\theta_2}.$$

But if S_1 and S_2 are the resistivities of the materials of the two conductors, and d_1 and d_2 their diameters, as determined by the micrometer gauge—

$$\frac{R_1}{R_2}=\frac{S_1 l_1}{d_1^2}\cdot\frac{d_2^2}{S_2 l_2}=\frac{\theta_1}{\theta_2}$$

$$\therefore \frac{S_1}{S_2}=\frac{\theta_1}{\theta_2}\cdot\frac{l_2 d_1^2}{l_1 d_2^2}$$

If it is desired to determine S_1 and S_2 in ohms for unit length and cross-section, a standard low resistance may be included in the

circuit and the deflection due to the p.d. across it compared with that for the given conductors.

By means of the *Kelvin bridge* a more accurate comparison of two low resistances may be made, but in this case the lengths l_1 and l_2 are adjusted until the resistances are equal, or in a ratio very nearly equal to unity.

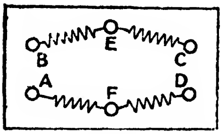


FIG. 92.

AD and BC (Fig. 92) are two resistance coils whose mid-points are at E and F, which are connected to the conductors under comparison as shown in Fig. 93. The current i_g in the galvanometer is zero when $\frac{r_2}{r_1}=m=n$. The conducting rods AB and CD (Fig. 93) are connected in series and a current passed through them. This current may be fairly strong, the rods having low resistance.

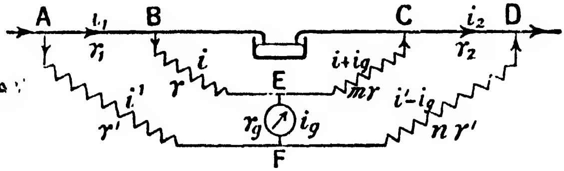


FIG. 93.

Let the resistance coil BEC be divided at E in the ratio $r : mr$, and the coil AFD in the ratio $r' : nr'$. The currents indicated in the diagram are consistent with Kirchhoff's first law (p. 65). Also the current in the rods between B and C is $i_1-i=i_2-(i+i_g)$, from which $i_2-i_1=i_g$.

Applying Kirchhoff's second law (p. 65) to the mesh ABEF, we have,

$$i_1r_1+ir-i_gr_g-i'r'=0,$$

and from the mesh CDFE,

$$i_2r_2-(i'-i_g)nr'+i_gr_g+(i+i_g)mr=0.$$

In making a measurement, one of the points A, B, C or D is moved along the rod until it is seen that the galvanometer deflection is zero, that is, $i_g=0$.

With this condition, the simultaneous equations reduce to,

$$i_1r_1+ir-i'r'=0$$
$$i_2r_2+mir-ni'r'=0,$$

and from above we see that when $i_g=0$, then $i_2=i_1$.

$$\therefore\ i_1r_1+ir-i'r'=0$$
$$i_1r_2+mir-ni'r'=0,$$

resistance r (Fig. 98), the current flowing in it will be $\frac{E}{R+r}$, where R is the resistance of the cell. Further, this current flowing in the conductor r means that a p.d. of $\frac{E}{R+r} \cdot r$ exists between the ends of the conductor, and this is proportional to l_2, the length of potentiometer wire which will now produce a balance. Calling this p.d. e, we have—

$$\frac{E}{R+r} \cdot r = e,$$

$$\frac{R+r}{r} = \frac{E}{e} = \frac{l_1}{l_2},$$

from which R may be found.

It is an interesting experiment to take a cell of the Leclanché type and determine its internal resistance a number of times,

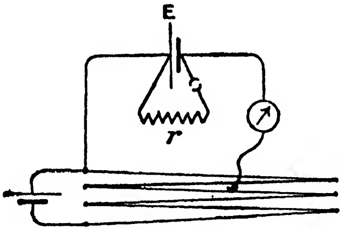


FIG. 98.

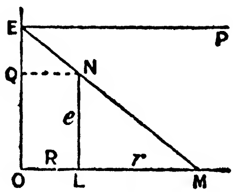


FIG. 99.

using values of r equal to 100, 50, 20, 10, 5 and 1 ohms respectively. It will be seen that the resistance of the cell varies with the current in it.

The reason for the method may be made a little plainer by drawing the E.M.F.—Resistance diagram for the circuit. Let OE (Fig. 99) be the E.M.F. of the cell; then, if the resistances of the different parts of the circuit be plotted along OM so that OL$=$R, LM$=r$, and the last point M be joined to E,

$$\tan \text{EMO} = \frac{E}{R+r} = \text{current};$$

further, we see that so long as Ohm's law holds good, the line, such as ENM, drawn upon the E.M.F.—Resistance diagram must be a straight line, since the current is the same at each cross-section of a circuit, and is represented on the diagram by the slope of the line. When the circuit is broken, the resistance r is infinite, and the point M moves to infinity; that is, the curve becomes the horizontal straight line EP. In the potentiometer

experiment, OE is measured in the first case (l_1), and LN in the second case (l_2), and we see from the figure that—

$$\frac{OE}{LN}=\frac{R+r}{r}.=\frac{l_1}{l_2}.$$

Rayleigh Potentiometer.—A potentiometer method in which two similar resistance boxes are employed in place of the stretched wire, was used by the late Lord Rayleigh in his work on standard cells (Fig. 100). The boxes are joined in series with a cell, to maintain steady current, and the plugs from one box are taken out. Any alteration in the resistances is made by transferring plugs from one box to the corresponding resistance gap of the other box, so that the total resistance in the two boxes remains constant, and the current therefore does not change. The E.M.F. of the cell E is then proportional to the resistance in the box A when the galvanometer indicates a balance.

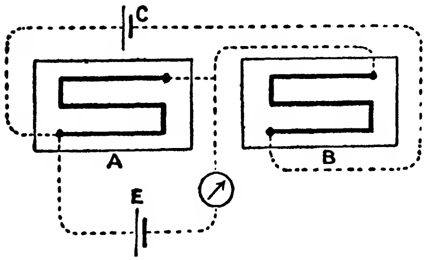


FIG. 100.

Crompton Potentiometer.—It will have been noticed that the smaller the current in the potentiometer wire, the greater the length of wire for a given p.d. and the more sensitive is the arrangement. For measuring very small E.M.F.'s, such as we have in the case of thermoelectromotive forces, which may be only a few millivolts, the sensitiveness may be sufficiently increased by including a resistance in the wire circuit, to reduce the current. But this decreases the range of the instrument, unless the added resistance has a value which is known in terms of the length of the potentiometer wire, in which case it may be included in the balancing part of the circuit when required. Thus if the balancing wire AB (Fig. 101) be 100 cm. long, and the resistances AC, CD, DE and EF are each equal to the resistance of AB, then, when the balance is attained for the point, say 35, upon AB, this corresponds to a p.d. proportional to 35 if the other contact is at A; but to 235 if at D, and 435 if at F, etc. This method of increasing the sensitiveness without sacrificing the range is used in the *Crompton potentiometer*, Fig. 102. *ab* is the balance wire, and at *c* there are fourteen coils, each of which has a resistance equal

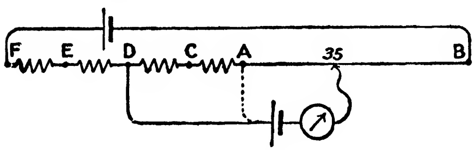


FIG. 101.

to the whole of *ab*, and so arranged that the potential terminal from the source under measurement may be connected directly to *a* or to the junction between any of the fourteen coils, by means of the rotating arm. *d* is a key which enables any of the external sources of E.M.F. to be compared, these being joined to A, B, C etc., to be rapidly and easily brought into action. Two adjustable resistances, *e* for rough and *f* for fine adjustment, enable the current to be varied, and if a standard Clark cell be connected to A, the temperature being 15° C., the E.M.F. is known to be 1·4328 volts, so that the whole 14

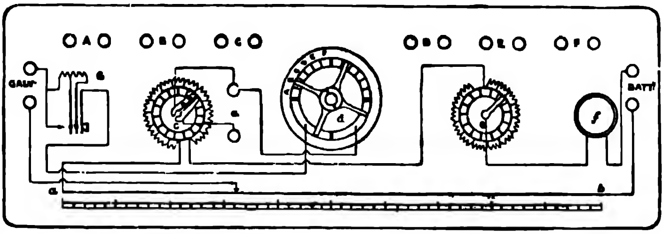


Fig. 102.

coils at *c* are switched in, and the contact set at the point 32·8 upon the wire, the current may be varied by these resistances until a balance is attained. Every unit of bridge wire will then correspond to $\frac{1}{1000}$ volt, and in this way the scale has been made to give readings directly in millivolts. This adjustment having been made, any other source connected to B may be measured directly in millivolts. Since the smallest scale division is one-tenth of a unit, and a balance may be made to say half a small division, electromotive forces down to $\frac{1}{20000}$ volt may be measured.

If the electromotive force being measured be the difference of potential between the ends of a standard resistance of $\frac{1}{100}$ or $\frac{1}{1000}$ ohm carrying a current, a unit on the slide wire will correspond to 0·1 ampere or 1 ampere respectively. Hence the usefulness of the instrument in conjunction with a standard resistance for current measurement.

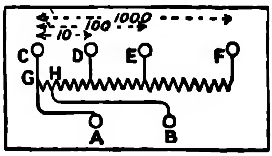


Fig. 103.

For measurement of large p.d.'s a volt box must be used. This is a high resistance, to the ends of which the p.d. to be measured is applied, the potentiometer terminals being joined to two points of the high resistance, separated by a small resistance which is some convenient fraction of the whole. Thus in Fig. 103 the resistance between G and H is $\frac{1}{1000}$ of that between C and F, $\frac{1}{100}$ of that between C and E, and $\frac{1}{10}$ of that between C and D.

If then a p.d. of say 120 volts exists between C and E, that between A and B is 1·2 volts, and on measuring this in the ordinary way by connecting A and B to the potential terminals of potentiometer, the observed value must be multiplied by 100. When C and D are used, the observed p.d. between A and B must be multiplied by 10, and when C and F are used, by 1000. Voltages less than 1·5 volts are measured directly without the aid of the volt box.

Direct-reading Resistance Meters.—Several types of instrument give scale readings of resistance. Some contain dry batteries from which current is passed through any resistance joined between the terminals of the instrument. If the unknown resistance is X and the whole internal circuit (including the movement of the meter) has a resistance R, then the current is $I = E/(R+X)$, E being the electromotive force of the battery. The current falls as X rises and becomes zero for X infinite. The resistance scale therefore reads in reverse direction to any current or voltage scale with which the instrument may be provided. Since the E.M.F. of the battery may vary with age and use, and so may its internal resistance, arrangement is made, *e.g.* by variation of subsidiary resistances in the circuit, to adjust the pointer to the zero-resistance mark when a short-circuit is applied to the terminals.

Instruments of this kind have very uneven scales for resistance. An alternative arrangement, especially suitable for measuring low resistances, is to have the moving coil shunted across the unknown resistance, with a large series resistance to keep the current through X nearly constant. The deflection, which is determined by the potential drop across X, is thus nearly proportional to X.

The current for these meters may be derived from a hand-driven generator, the E.M.F. of which is kept constant by a mechanical device which prevents the armature from rotating above a certain speed. Another way of ensuring that the result is independent of variations in supply voltage is to use a pair of coils rigidly connected, usually at right angles, and pivoted as freely as may be in the space between the magnet poles of the meter. One coil is in series with a fixed resistance and carries a current proportional to the E.M.F. of the generator while the other is in series with the unknown resistance, and the generator feeds both. The coil-system turns until the resultant magnetic moment of the pair lies along the magnetic field. The rotation required to attain this condition depends on the ratio of the currents in the two coils and hence on the ratio of the potential difference across X to the current in it. Thus the scale over which the pointer moves may be marked in ohms. This form of instrument is particularly useful for testing insulation resistances and other high resistances.

CHAPTER IV

ELECTROSTATICS

Early Electrical Experiments.—The earliest electrical experiment of which we have any record is that of the attraction of light bodies by a piece of amber that had previously been rubbed. This experiment is of unknown antiquity, but William Gilbert (1540–1603), when investigating this and other allied phenomena, introduced the word " electricity," from the Greek word ἤλεκτρον, signifying " amber." Other substances exhibit similar properties, and if, for example, a piece of ebonite be rubbed with a piece of dry fur, then on separating them they will be seen to attract each other ; on bringing the ebonite near the fur the individual hairs of the latter will bend towards the ebonite. Simple experiments show that an ebonite rod rubbed with fur will repel another one similarly treated. Also a glass rod rubbed with silk repels a similar glass rod. But the ebonite rod attracts the glass rod. This is usually explained by saying that there are two kinds of electricity, that on the glass rubbed with silk being called *positive* electricity and that on the ebonite rubbed with fur *negative* electricity. Also like kinds of electricity repel each other and unlike kinds attract each other.

About a century after Gilbert's time, it was found that all substances taken in pairs became oppositely electrified when rubbed together, but in the case of conductors, the electrification disappears as soon as the bodies are separated. If, however, a metal rod be held by an insulating handle it can easily be electrified by rubbing with fur or silk.

The attraction of the charged ebonite or glass for light bodies is easily understood if these bodies are conducting. On referring to Fig. 107 it is seen that the charged rod causes an opposite charge to accumulate on the near side of the light body. There is consequently an attraction between the charged rod and the light body.

Theories of Electrification.—Several theories have been put forward to account for the phenomenon of electrification. The earliest theory was that two weightless fluids exist, positive and negative electricity, and the preponderance of either of them on a body determines the sign of the charge on it. There was also a one-fluid theory, for it was pointed out that if a body is un-

charged when there is a normal amount of this fluid on it, then an excess of it meant a charge of one kind and a deficiency a charge of the opposite kind.

Maxwell, following Faraday, suggested a model rather than a theory. Thus, positive electrification he considered to be due to a displacement of something outwards from a conductor and negative electrification a displacement inwards. This displacement is then continuous in the hypothetical medium separating the conductors, and its rate of change is equivalent to an electric current.

It had been shown by Faraday that a voltaic current is in all respects equivalent to the motion of the electricity produced by friction.

The investigations of the last half-century have shown that electricity is not a fluid, in the ordinary sense, but that it consists of particles of definite amount. The *electron* (Chap. XIV) is a definite unit of electricity and is, so far as we know at present, indivisible, and its charge is of the kind which has been called *negative*. At the same time there are corresponding positive charges, *positrons* (Chap. XVI), exactly like electrons but of opposite sign. Then there are heavier bodies having the charge of the positron, which constitute most of the mass of atoms of matter. These are called *protons*. A proton and an electron together constitute an atom of hydrogen. Also there are *neutrons* which have the mass of a proton but no electric charge. All atoms of matter are built up of these bodies, but whether a proton is a combination of a neutron and a positron, or whether the neutron is a combination of a proton and an electron is not yet known.

Our present knowledge gives us a clear explanation of many phenomena which were obscure on the older theories. For example, a conductor of electricity is a substance in which electrons are readily detached from the atoms and exist as a cloud or gas in the interspaces between the atoms. The presence of an electric field causes a drift of the electrons through the conductor, which we recognise as an electric current. From the laws of electrodynamics it is then concluded that an electron in motion has both an electric and a magnetic field. When at rest it has an electrostatic field only. Thus, if an electron is detached from a neutral atom, the atom has then a positive charge. The atoms of a solid body are so close together that they are held in position in various structures (Chap. XIV).

In a non-conductor or dielectric the electrons are not detachable from the atoms, so that there is no possibility of a continuous current. There is, of course, no sharp division between conductors and non-conductors. This consideration throws light upon the behaviour of a dielectric in an electric field.

The theory adopted for the explanation of electrification and conduction does not affect the principal problems of electrostatics, and, except for the consideration of dielectrics, the discussion of the electron theory will be deferred to a later chapter.

The Electroscope.—It has been seen that an electric charge may be detected by its ability to attract light bodies. A pith ball, suspended by a silk fibre, serves as a suitable light body. A modification consists in suspending two pith balls side by side by means of conducting threads. The charges induced upon them cause repulsion and the balls stand apart.

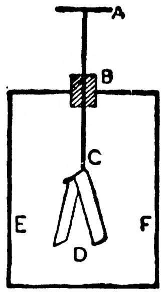


FIG. 104.

If two strips of thin foil, gold or aluminium, replace the pith balls a more convenient and sensitive apparatus, called an electroscope, results. In Fig. 104 the leaves or strips of foil are shown at D and are supported by the wire C which passes through a block of paraffin wax or sulphur B. Strips of tinfoil are pasted at E and F opposite the leaves and at the back and front are glass windows. On giving a charge to the cap A it will spread over A, C and D and the leaves will stand apart.

On giving a charge of, say, positive electricity to the leaves, it will be found, on bringing a positively charged body near A, that the divergence of the leaves increases. On the other hand, a negatively charged body brought near A will cause a lessening of the divergence.

Faraday's Ice-pail Experiments.—By means of the electroscope, Faraday was enabled to establish several important laws of electric phenomena. He used an ice pail, which gave its name to the experiments, supported on an insulating stand and connected with an electroscope (Fig. 105). On lowering a charged conducting body by means of an insulating silk thread, into the ice-pail, the leaves of the electroscope diverge more and more until the charged body is well inside the pail, and then movement of the body about inside the pail produces no further change. If the charged body be removed without having touched the pail, the leaves collapse completely, but if before removal it be allowed to touch the pail on the inside, no alteration in the divergence of the leaves is produced on touching, or when the body is removed. Its charge has therefore passed completely to the outside of the pail. It is, therefore, reasonable to suppose that the divergence of the leaves depends upon the amount of charge on the body lowered into the pail, and equal charges may be compared

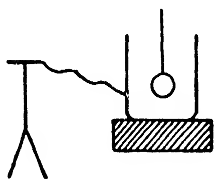


FIG. 105.

by lowering them in turn into the pail and noting that the divergence is the same for each. Further, if equal charges of opposite kinds be obtained upon two separate bodies, that is, two charges of opposite kinds that would each produce the same divergence when used alone, then when placed inside together, whether the bodies touch or not, there will be no divergence of the leaves, showing that the effect of adding equal and opposite charges is to produce a neutral condition.

A similar experiment may be made, to show that the amounts of electricity produced in two bodies in the act of rubbing are equal, and that they are of opposite kind. An ebonite cylinder, C (Fig. 106), is made to fit loosely into a hollow cylinder, F, lined with fur, and the two are placed inside the ice-pail. If now C be given a few turns by means of the handle, there is no divergence of the leaves of the electroscope, but on removing C the leaves diverge on account of the charge on F. On removing F and putting C in the pail, the leaves diverge to the same extent as for F, but when F and C are both in the pail together the divergence is always zero, showing that the electric charges on the two are equal and of opposite sign.

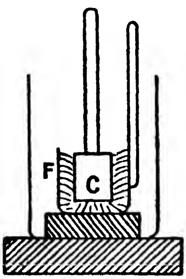


FIG. 106.

Potential Gradient due to Charge.—The student has probably noticed already the similarity in the effect of charge upon charge to that of a magnetic pole upon a second pole. The exact analogy between the laws will be treated later, but, as on p. 13 the repulsion between N poles implies that there is a magnetic potential which decreases as the distance from a N pole increases, so in the neighbourhood of a positive electric charge there is an electric potential which diminishes as the distance from the charge increases. Also a N pole experiences a force urging it down the gradient of magnetic potential, so a positive charge experiences a force urging it down the gradient of electric potential. It must always be borne in mind that the analogy is only mathematical. There is no such thing as a conductor of magnetic pole, in fact the pole itself is a fiction, although a very useful one. That a charge produces a potential gradient in its neighbourhood may be shown by employing an insulated conductor AB (Fig. 107) which is divided into two parts, the parts being in contact. On bringing a positively charged body C near to B a potential gradient exists, B being at a higher potential than A. Hence a current takes place

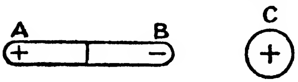


FIG. 107.

appearance will be seen in Fig. 111. When the machine is running, the positively and negatively charged conductors B and B′ come opposite to C and C′ at the instant that the latter are connected together by two light brushes carried by a wire. Then a potential difference exists between C and C′, and a current flows from C to C′. Since these were uncharged before, C will now be negatively and C′ positively charged, and the charges, moving on as the plates revolve, play a similar part in charging the outer sectors at B″ and B‴. The result of the two processes is that positively charged sectors B are continually being brought into

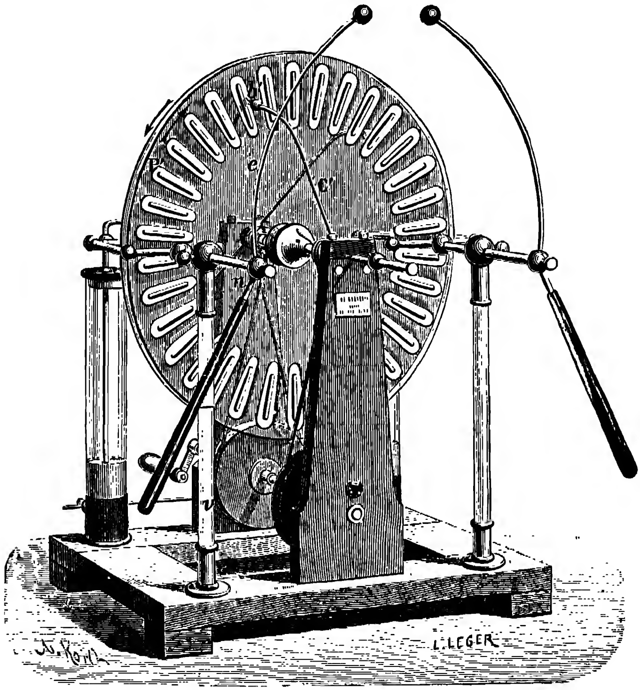


FIG. 111.

the neighbourhood of the collector E, and C into the neighbourhood of E′, while the negatively charged sectors are brought to F and F′. From the collectors E and E′ the positive charge passes to the conductor P, and from F and F′ the negative charge goes to Q. The method by which the collection of the charge by E and F takes place may be noted. As the positive charge approaches E it induces charges upon E, as in the experiment described on p. 108 (Fig. 107), but the negative charge, being situated on the points nearest to the sectors, passes readily from them on to the sector and neutralises the positive charge there,

the sector being thereby discharged, and a corresponding positive charge remaining upon the conducting system EP. A similar process occurs at F, but in this case it is a negative charge that remains on FQ. The action of a point in facilitating a discharge is described on p. 136.

Force between Charges.—From analogy with the gravitational force between two masses, it would naturally be suggested that the force between two small electrical charges would vary inversely as the square of their distance apart. By means of his *torsion balance,* Coulomb established roughly the truth of this law. He balanced the force between the charges on two gilt pith balls, one of which was fixed, and the other on the end of a light rod suspended by a fine silver wire, against the force produced by the measured twist in the wire. On halving the distance between the balls the necessary twist in the wire was increased four times, and so on. Further, by removing the fixed ball and sharing its charge with an equal one (the charge on the original ball, by the principle of symmetry, being supposed to be halved) and replacing the ball, it was found that the force as indicated by the torsion balance was halved. Coulomb concluded that the force between two charges may be represented by the expression $\frac{q_1 q_2}{d^2}$, where q_1 and q_2 are the magnitudes of the charges, and d their distance apart, the positive sign indicating that the force between the charges is of the nature of a repulsion.

From this relation we may define the *unit of electrical charge as that which, situated one centimetre from an equal charge in empty space, or what is nearly the same thing, in air, will repel or attract it with a force of one dyne;* and adopting this new electrostatic unit of quantity and measuring q_1 and q_2 in terms of it, we have—

$$\text{Force between charges} = \frac{q_1 q_2}{d^2} \text{ dynes.}$$

By analogy with the magnetic case (p. 3), we may define the *strength of electric field,* or *electric intensity,* or *electric force* at a point as the force in dynes, which would act on a unit positive charge if placed at that point, and we see then that the electric intensity at a distance d centimetres from a charge q is $\frac{q}{d^2}$, and that the force on any charge at a point at which the electric intensity is E, is equal to Eq.

Proof of Inverse Square Law.—The proof of the relation $F \propto \frac{q_1 q_2}{d^2}$ by means of Coulomb's torsion balance is not very satisfactory, because the charges are not situated at points, but are distributed

over metallic spheres, and although this would not matter if the charges were uniformly distributed over the spheres, this condition cannot be fulfilled, since the presence of each charged sphere would disturb the distribution of charge upon the other. Also charges will be produced upon the case of the instrument, and upon all other conductors in the neighbourhood, and further, the holders of the charged balls are not perfect insulators, so that the charges will gradually leak away; and finally the amount of torsion and the distance apart of the balls cannot be measured very accurately. We have, therefore, to fall back upon indirect methods of proof, that is, to calculate certain results on the assumption of the truth of the law, and then put the results to the test of experiment.

The following proof is due to Cavendish; and at a later date Maxwell re-performed the experiment and succeeded in showing that the inverse square law is certainly very near the truth. To find the strength of electrical field at a point P (Fig. 112), situated within a charged spherical conductor, draw through P a cone having its vertex at P, and whose solid angle $d\omega$ is very small. This cone cuts the sphere in two small areas, ds and ds_1, which may easily be found from the distances r and r_1 of P from ds and ds_1 respectively. The area of the right section of the cone at ds is $r^2 d\omega$, and this makes angle α with ds, $\therefore\ ds = \dfrac{r^2 d\omega}{\cos\alpha}$, and similarly $ds_1 = \dfrac{r_1{}^2 d\omega}{\cos\alpha}$. If the sphere is symmetrically situated with respect to neighbouring conductors, the charge upon it will be uniformly distributed. Let the amount of charge on each unit of area of surface be σ; then the amount on ds is $\dfrac{r^2\sigma d\omega}{\cos\alpha}$, and that upon ds_1, $\dfrac{r_1{}^2\sigma d\omega}{\cos\alpha}$. If, then, the strength of field due to a charge varies inversely as the nth power of the distance, the field at P due to the charge on ds will be $\dfrac{r^2\sigma d\omega}{r^n\cos\alpha}$, and due to that on ds_1, $\dfrac{r_1{}^2\sigma d\omega}{r_1{}^n\cos\alpha}$. These are obviously equal when $n=2$, and since they are oppositely directed, the resultant field at P due to the charges on ds and ds_1 is zero. The whole sphere may be divided by cones into pairs of surfaces in the same manner, and consequently the electrical intensity at P due to the whole charge on the sphere is zero.

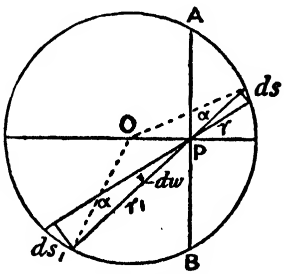


FIG. 112.

If $n>2$ the component of intensity due to ds is greater than that due to ds_1, since $\frac{1}{r^{n-2}}$ will then be greater than $\frac{1}{r_1{}^{n-2}}$, r_1 being greater than r, and all the elements on the same side of the plane APB (Fig. 112) as ds, give rise to components at P greater than those due to the corresponding elements on the same side of the plane as ds_1, since for all these pairs $r_1>r$, and if the charge on the sphere be of positive electricity, there will be a resultant field towards the centre of the sphere. On the other hand, if $n<2$ it follows in a similar manner that there will be a resultant field which will be directed outwards from the centre.

Cavendish, and at a later date Maxwell, supported a sphere, A, inside a second sphere, B (Fig. 113), so that the two are independently insulated, except when connection is made between them by the hinged wire at the top. B is first positively charged ; then A is connected to B by means of the wire, and the connection is then broken so that A is again insulated. From the reasoning given above, A would be positively charged if $n>2$ and negatively charged if $n<2$, and uncharged if $n=2$. Cavendish, using a pith-ball electroscope, could not detect any charge upon A, and from the result of his experiment concluded that n must certainly be within one per cent. of the value 2. Later Faraday, with a gold leaf, failed to detect any charge within a closed conductor, and in 1870 Maxwell re-performed Cavendish's experiment, using a quadrant electrometer to detect the charge, and again failed to find any. From a measurement of the smallest charge upon A that could be detected by the electrometer, Maxwell concluded that n cannot differ from 2 by more than $\frac{1}{21600}$. Plimpton and Lawton,[1] applying a low-frequency high potential to the outer sphere and studying potentials developed across a high resistance joining the two concentric spheres, estimated the value 2 to be true to better than 1 in 10^9.

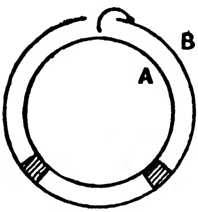


FIG. 113.

Dielectrics.—Until the discovery of the electron (Chap. XIV) there was no detailed knowledge of the structure of the atom. As a consequence electric phenomena were treated as though material substances were homogeneous in structure. Each dielectric, including empty space, was supposed to possess some property, designated by a number known as the *dielectric constant* or *specific inductive capacity* k. Thus the capacity of a condenser was found, by Faraday, to depend upon the dielectric between the conducting plates, the capacity being increased k times by replacing air by a dielectric of constant k. Also the

[1] S. J. Plimpton and W. E. Lawton, *Phys. Rev.*, **50**, p. 1066. 1936.

force between two charges q_1 and q_2 is $\frac{q_1 r_2}{kr^2}$ when they are situated in a dielectric of constant k.

It is now known that all atoms possess electric charges in the form of electrons, protons or positrons (Chap. XVI). We are not concerned here with the structure of the atom. The only property we are interested in, at the moment, is that the electron is detachable from the atom in the case of conducting substances and the fact that in dielectrics the electron cannot be liberated from the atom. Whatever the configuration of the charges in the neutral atom may be, they do not produce any external field, but the effect of an electric field upon it will be to displace positive charge in the direction of the field and negative charge in the opposite direction. These displacements are limited by the forces that bind the charges to the atom. Each atom is then a *dipole* or *electric doublet* and has an *electric moment* exactly analogous to the magnetic moment of a small magnet (p. 5) and will produce an electric field in its neighbourhood. If x is the distance separating the effective charges $+q$ and $-q$ the electric moment is qx, the potential due to it is $\frac{qx \cos \theta}{r^2}$ (p. 16) and the field is $\frac{qx}{r^3}\sqrt{1+3 \cos^2 \theta}$ (p. 19). Of course there may be various charges displaced within the atom, but each pair will have a moment and qx is taken as the vector sum, or resultant, of all these moments.

Electric Polarisation.—Only isotropic dielectrics will be dealt with here, that is, dielectrics which have identical electrical properties in all directions. Consider a rectangular block of the dielectric situated in an electric field E parallel to the edge AB of the block (Fig. 114). Owing to the atoms having electric moments, the electric moment of the block will be the resultant of the electric moments of the atoms contained in it. If l is the length AB and a the area of the face AC, la is the volume, and if P is the electric moment per unit volume, considered as uniform, Pla is the moment of the whole. The quantity P is called the *polarisation* of the medium and it may be taken as established by experiment that P is proportional to E. This may be represented by the relation $P=\epsilon E$, where ϵ is a constant quantity for each dielectric, but varying from one to another, being zero for a vacuum. Again, if AC represents the boundary of the dielectric in zero field, the presence of the resultant field E causes a positive charge to appear over AC and a negative

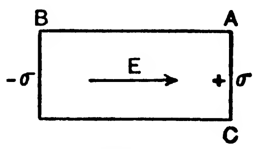


FIG. 114.

charge over the opposite end. If σ is the surface density of this charge, $+\sigma a$ is the charge on the face AC and $-\sigma a$ the charge on the opposite face. The electric moment of the block is thus σal, and equating this to Pal, it is seen that $P=\sigma$.

Even though the material may be continued beyond AC, P, or σ, is still the charge per unit area appearing on AC. If FG (Fig. 115) is the layer of atoms in this continuation, the sheet of charge $+\sigma$ on AC is faced by a sheet $-\sigma$ on FG. There is consequently a field of intensity $2\pi\sigma$ due to AC and $2\pi\sigma$ due to FG (p. 126) between AC and FG, and as these are in the same direction the resultant of the two is $4\pi\sigma$ or 4πP.

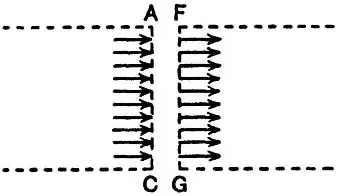


FIG. 115.

If the dielectric were absent, the electric intensity would be E, but owing to the dielectric it has become, in the space between AC and FG, $E+4\pi P$. This quantity is called the *electrical induction* within the dielectric, and giving it the symbol ϕ,

$$\phi=E+4\pi P.$$

Effects of Polarisation.—(i) *Point charges.* It is now possible to account for all electrostatic effects in terms of the electric field which may exist in empty space, or what is nearly the same thing, in air, together with any field due to the polarisation of the dielectric. Thus, the force between two point charges, q_1 and q_2 situated a distance r cm. apart in empty space is $\frac{q_1q_2}{r^2}$ dynes. Now consider that q_1 and q_2 are situated in a dielectric. Being point charges, q_1 produces a field at q_2 and in addition a radial polarisation everywhere, designated by P. If q_1 is a positive charge (Fig. 116) the negative ends of the dipoles of the dielectric are turned towards it, and *vice versâ*. Imagine now a sphere of radius x to be drawn, with q_1 as centre. Let E be the resultant field at the surface of the sphere; then $P=\epsilon E$. If we imagine the dielectric within the sphere to be removed, there is a negative charge on the concave spherical surface so formed, and in order to calculate the resultant field outside the sphere, q_1 must be reduced by the amount $4\pi x^2P$. The resultant field E at the surface is therefore $\frac{q_1-4\pi x^2P}{x^2}$, and $P=\epsilon E=\epsilon\frac{q_1-4\pi x^2P}{x^2}$

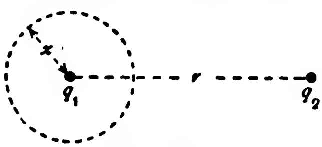


FIG. 116.

$$\therefore\ P=\frac{\epsilon q_1}{x^2(1+4\pi\epsilon)}$$

so that the charge to be subtracted from q_1 is $4\pi x^2 P = \frac{4\pi\epsilon q_1}{1+4\pi\epsilon}$ and the effective charge is

$$q_1 - \frac{4\pi\epsilon q_1}{1+4\pi\epsilon} = \frac{q_1}{1+4\pi\epsilon}.$$

Thus the force on q_2 is $\frac{q_1 q_2}{(1+4\pi\epsilon)r^2}$ instead of the value when there is no dielectric, $\frac{q_1 q_2}{r^2}$. It might be thought that the production of the cavity might invalidate the calculation. But it should be noticed that x may have any value, less than r, and may be diminished until the spherical surface contains the very last atoms surrounding q_1.

(ii) *Field near Plane Conductor.*—If A (Fig. 117) is a plane conductor facing a parallel plane conductor B, and having a surface density of charge $+\sigma$, there is a surface density $-\sigma$ upon B, and the field between them is $4\pi\sigma$ (p. 126). Now imagine a dielectric to be introduced and fill the space between A and B. If the resultant field in this space be E, the polarisation of the dielectric is ϵE, and this implies a surface density $-\epsilon E$ on the surface of the dielectric in contact with A and $+\epsilon E$ on the surface in contact with B. The effective surface densities are therefore $+(\sigma - \epsilon E)$ and $-(\sigma - \epsilon E)$ and the field E is therefore $4\pi(\sigma - \epsilon E)$.

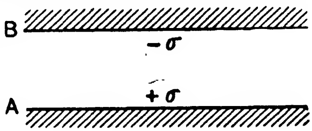


FIG. 117.

That is, $$E = 4\pi(\sigma - \epsilon E)$$

or, $$E = \frac{4\pi\sigma}{1+4\pi\epsilon}$$

instead of $4\pi\sigma$ in the absence of the dielectric.

Dielectric Constant.—In the above results the quantity $1+4\pi\epsilon$ appears in problems with dielectrics in place of unity when the problem refers to charges in a vacuum. The quantity $4\pi\epsilon$ is due to the polarisation of the dielectric. It is possible to continue to use this expression, but it is convenient to substitute one letter for it, that is, to write $1+4\pi\epsilon = k$. The name *dielectric constant* or *specific inductive capacity* is given to the quantity k, and the above results may then be written

$$\text{Force between charges} = \frac{q_1 q_2}{k r^2}$$

$$\text{Electric intensity near plane conductor} = \frac{4\pi\sigma}{k}.$$

The equation $k=1+4\pi\epsilon$ may be written $\mathrm{E}k=\mathrm{E}+4\pi\epsilon\mathrm{E}$ $=\mathrm{E}+4\pi\mathrm{P}=\phi$ . . . (p. 116). It leads then to a new definition of the quantity ϕ, or *electrical induction* which is now written $k\mathrm{E}$. It should be noted that many writers use the symbol D for $k\mathrm{E}$, but here the letter D has been retained for the electric displacement, as defined by Maxwell, and $\mathrm{D}=\frac{k\mathrm{E}}{4\pi}$ (p. 128).

In the expression for the force between charges in empty space, $\mathrm{F}=\frac{q_1q_2}{r^2}$ (p. 112), the charges were so chosen that the force should be in dynes. It is now seen that the medium in which the charges are situated plays a part in determining the force. This raises the question of the justification for making $k=1$ for empty space, which was tacitly done. As a matter of fact it is not possible to give a value to k for empty space. The electric properties of dielectrics are found to exist in empty space. The chief of these are the ability to support an electric field, and the inability to conduct electrically. For these reasons it is often desirable to introduce a quantity k_0 into our equations when they apply to charges situated in a vacuum. Thus $\mathrm{F}=\frac{q_1q_2}{k_0r^2}$ dynes is an expression of the fact that the only measurable effects of charges at rest are the forces they produce. Continuing on the same plan, $k\mathrm{E}=k_0\mathrm{E}+4\pi\epsilon\mathrm{E}$

or,
$$k=k_0+4\pi\epsilon.$$

It is possible then to define the specific inductive capacity of a dielectric as the quantity $\frac{k}{k_0}$ which is also equal to the ratio,

$$\frac{\text{force between two charges in vacuum}}{\text{force between similar charges in dielectric}}$$

$$=\frac{\text{capacity of condenser with the dielectric}}{\text{capacity of the condenser with vacuum between plates}}.$$

The last ratio is in agreement with the original definition due to Faraday (p. 114).

Potential.—Magnetic potential was defined on p. 13 ; electrical potential may be defined in a similar manner.

Potential is a quantity whose rate of variation in any direction is the electric intensity or force in that direction.

Thus $\mathrm{E}=-\frac{d\mathrm{V}}{dx}$, where V is electric potential. This may also be written in the form—

$$d\mathrm{V}=-\mathrm{E}dx.$$

Edx is the amount of work done, when a unit charge is moved through the infinitesimal distance dx. Hence if the unit charge be carried from a point a to a point b along a path whose direction everywhere coincides with that of E, the total work done is

$$-\int_a^b \mathrm{E}dx$$

But,
$$\int_a^b d\mathrm{V}=-\int_a^b \mathrm{E}dx$$

that is, $\mathrm{V}_b-\mathrm{V}_a$=work done in carrying unit charge from a to b.

We may, if we choose, define potential from this relation as a quantity the difference in whose values at two points is the amount of work which must be done in carrying a unit charge from one point to the other.

The signs are so chosen that a positive field is directed away from a positive charge, points nearer to it being at higher potential than those further away. Hence from the last relation we see that work is to be considered as negative when the charge is moving down the grade of potential, and therefore positive when up the grade of potential. Thus work is to be considered to be positive when it is done upon the charge by any external agency.

It is immaterial what path is followed by the charge in passing from a to b. The component of E along the path at a point such as P is E cos θ, and the work done for the element dl measured along the path is E cos $\theta \,.\, dl$.

Also,
$$\frac{d\mathrm{V}}{dl}=-\mathrm{E}\cos\theta, \text{ or, } d\mathrm{V}=-\mathrm{E}\cos\theta\,.\,dl$$

$$\mathrm{V}_a-\mathrm{V}_b=\int_a^b \mathrm{E}\cos\theta\,.\,dl.$$

The work done in passing along the path aPb (Fig. 118) is equal to that in traversing aQb, for if it were not there would be a balance of work done in describing the cyclical path aPbQa, so that a would have more than one potential. In purely electrostatic phenomena there is only one value of the potential at each point, and the truth of our proposition depends upon this fact, although it is by no means true for all possible cases; in fact, we shall see later that the integral $\int$ F cos $\theta \,.\, dl$ for a closed path, which we call the line integral of the field F round the circuital path, measures in some cases the flux of some quantity through the plane of the circuit. In these cases, however, the phenomenon is not purely statical. It will be discussed more fully in Chapter VIII.

FIG. 118.

The definition of potential given above is consistent with that

for potential difference on p. 54. For the current being the amount of electricity which passes through a given section of the conductor per second—

$$i=\frac{q}{t} \quad \text{or,} \quad q=it.$$

Hence the work done by the current in the section of the conductor between the points a and b for a charge q to pass, is—

$$q(V_a-V_b)=i(V_a-V_b)t.$$

Thus the work done per second is $i(V_a-V_b)$, that is, the product of current and p.d. It must be remembered, however, that different units of quantity of electricity are used in the two problems. For convenience in studying electrostatic effects, the electrostatic unit of quantity defined on p. 112, which is derived from the force between charges, is employed, while that which is derived from the force between magnetic poles, by way of the magnetic field, and electric current is called the electromagnetic unit of charge. The relation between the two units is a very important one, and will be discussed in Chapter XII.

The *potential due to a charge* $+q$ measured in electrostatic units, follows from our knowledge of the electric intensity due to $+q$.

FIG. 119.

Thus, if a, x and b are the distances of these points from $+q$ (Fig. 119), then at x—

$$E=\frac{q}{x^2}, \qquad \frac{dV}{dx}=-\frac{q}{x^2}, \qquad dV=-\frac{q}{x^2}dx.$$

Thus the work done in carrying unit positive charge from b to a in opposition to the force due to q is the excess in potential at a over that at b, and thus,

$$V_a-V_b=-q\int_b^a\frac{dx}{x^2}=q\left[\frac{1}{x}\right]_b^a=\frac{q}{a}-\frac{q}{b}.$$

If the unit charge were carried from a to b instead of from b to a, work would be done by the field upon the charge, and V_a-V_b would be negative. Thus a has a higher potential than b in accordance with our convention (p. 119).

Potential is only measurable by its differences, and therefore there is no absolute zero of potential. It is convenient, however, to consider all points at infinity to be at zero potential, since the field at an infinite distance from our electric charges is zero. Thus if the point b be situated at infinity, $\frac{1}{b}=0$ and potential at a is $\frac{q}{a}$.

Equipotential Surfaces.—An *equipotential surface* is one drawn through a system of points which are at the same potential. In Fig. 120 the equipotential surfaces due to a charge +20 units are spheres which appear in section as circles in the diagram. Those due to the second charge of 20 units are also shown, and in the left-hand upper half of the figure the resultant equipotential surfaces are shown as dotted lines, for the case in which this second charge is negative. The potential at any point is the algebraic sum of the potentials due to the two charges. In the right-hand upper part, the equipotential surfaces are drawn for the case in which both charges are positive.

The electric intensity at any point is at right angles to the equipotential surface passing through the point, for if this were

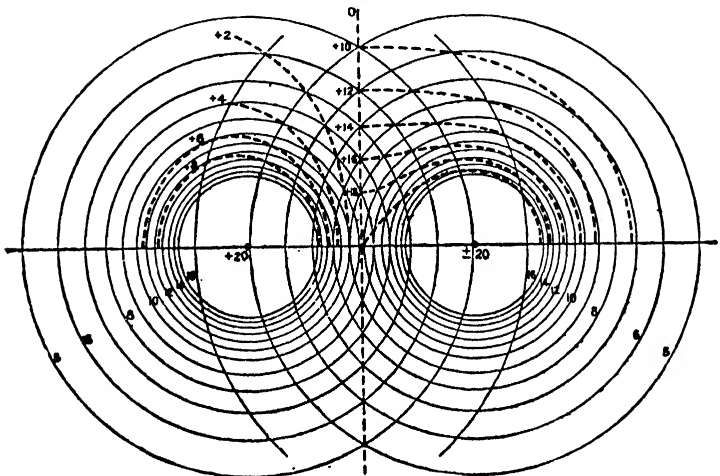


Fig. 120.

not the case, it would have a component along the surface, which is only another way of saying that the potential varies as we pass from point to point along the surface. The surface would then not be one of equipotential. Thus if the equipotential surfaces for a given field be known, the direction of the intensity at every point may be found by drawing a system of lines that cut the equipotential surfaces everywhere at right angles. Lines drawn in this way would be curves of a similar shape to those in Fig. 9 (i) when the charges have opposite signs, and to those in Fig. 9 (ii) when both charges have the same sign.

The surface of a conductor, upon which any charges which may be present are at rest, is necessarily an equipotential surface, since if this were not the case there would be an electric intensity other than zero directed along the surface, in which case a current

would flow. It follows that in any electrostatic problem we may imagine a conducting surface to coincide with any equipotential surface, on giving it the requisite potential, without in any way changing the conditions of the problem. This device is often of great convenience, as we shall see later.

Energy of Charge.—The process of placing a charge upon a conductor necessitates the expenditure of a certain amount of energy, which may be derived from a variety of sources. In charging a body by friction, equal amounts of positive and negative electricity are in contact until the bodies on which they reside are separated, and the action of separation requires a mechanical force to overcome the attraction between the charges. Similarly, work must be done in removing the charged metal plate from the oppositely charged ebonite sheet of the electrophorus; and the plates of the Wimshurst machine require the expenditure of work in turning them in opposition to the attraction between the oppositely charged conductors B and C (Fig. 110). The work done is stored up as potential energy upon the charged body, and supplies the energy necessary to drive the current when the conductor is discharged.

The energy possessed by the body on account of the charge residing upon it may be expressed in terms of the amount of charge and the potential of the body. Thus if v is the potential of the body, this represents the amount of work necessary to bring a unit charge from a point at zero potential and place it on the body; so that, to add the infinitesimal charge dq, the work necessary is vdq. But we have seen that at every point in the neighbourhood of a charge, the potential due to it is proportional to the charge;

$$\therefore\ v=aq, \text{ where } a \text{ is some constant,}$$

$$\text{and work for increase of charge } dq \text{ is } aqdq$$

$$\therefore\ \text{work for finite charge Q is } \int_0^Q aqdq=\tfrac{1}{2}aQ^2.$$

But the potential for charge Q is—

$$aQ=\text{say V},$$

$$\therefore\ \text{energy}=\tfrac{1}{2}QV.$$

If the body be discharged by a conductor, the heat produced in the conductor is therefore $\frac{1}{2}QV$, provided that none of it is used in producing chemical action, light, sound etc.

The inverse of the quantity a, we shall see in the next chapter, is called the *capacity of the conductor*, and the energy of the charge may be expressed in terms of it. Thus if $C=\frac{1}{a}$,

$$\text{Energy of charge} = \tfrac{1}{2}\frac{Q^2}{C},$$

or since $Q = CV$,

$$\text{Energy} = \tfrac{1}{2}QV = \tfrac{1}{2}CV^2.$$

Theorem of Gauss.—Since an electric charge is surrounded by an electric field whose intensity has a definite value at every point, we should expect that a knowledge of the intensity at every point of a closed surface surrounding the charge would enable us to determine the charge. The case is analogous to that of the uniform extrusion of a fluid from a point source situated within the fluid, for we can calculate the rate of extrusion by finding the total volume of liquid which crosses a closed surface, surrounding the point, in unit time. In fact the two problems are mathematically very similar, the solution for the case of an incompressible fluid being of the same form as that for an electric field.

In the electrical problem, the quantity to be evaluated for the whole closed surface is called the *electric flux*. It may be defined as the product of the normal component of electrical induction and the area over which the induction may be considered constant. It is thus $\phi \cos \theta ds$, where ϕ is the induction, θ the angle it makes with the normal to the surface at P (Fig. 121) and ds a small element of surface at P. For empty space

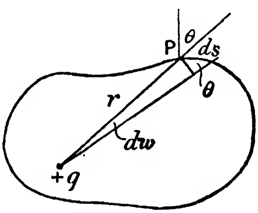


FIG. 121.

$\phi = k_0 E$ (p. 118) and $E = \dfrac{q}{k_0 r^2}$, where the charge is $+q$; and ϕ and E are outwardly directed through the surface. Thus $\phi \cos \theta \,.\, ds = k_0 \dfrac{q}{k_0 r^2} \cos \theta ds = q \,.\, \dfrac{\cos \theta \,.\, ds}{r^2}$ is the electric flux through ds. Now $\dfrac{\cos \theta \,.\, ds}{r^2}$ is the solid angle $d\omega$ subtended at q by the surface ds.

$$\therefore \text{ flux for element } ds = q d\omega$$

and total flux for the whole closed surface is

$$\int q d\omega = q \int d\omega = 4\pi q,$$

since q is constant, and the solid angle $\int d\omega$ subtended by the whole closed surface is 4π.

When a dielectric other than empty space is situated at ds, the flux is $k \,.\, \dfrac{q}{kr^2} \cos \theta ds = q \,.\, \dfrac{\cos \theta ds}{r^2}$, and is the same as for empty space.

If there are more charges than one within the surface, each charge q contributes an amount $4\pi q$ to the normal flux over the whole surface, and if q is positive the flux is directed outwards, if negative, inwards, so that Gauss's theorem may be stated—

The total electric flux over a closed surface is 4π times the total amount of charge within the surface ($\Sigma 4\pi q$).

If there is no resultant charge within the surface, $\Sigma 4\pi q=0$, and therefore the total electric flux over it is zero, and *vice versâ.*

We can easily see that if a charge be situated outside the surface, it does not contribute anything to the total electric flux over the surface. For at P (Fig. 122) the electric flux for an element of surface is $-qd\omega$, the negative sign being taken when the intensity is directed inwards, and at Q it is $+qd\omega$, the direction being outwards, so that these two elements together would give a resultant of zero for the electric flux. Similarly for any other cone drawn through q to cut the surface, and the total electric flux is therefore zero.

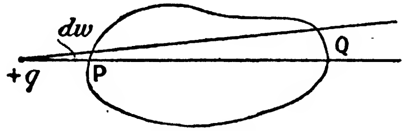


FIG. 122.

Electric Intensity near Charged Sphere.—Many useful problems may be very simply solved by applying Gauss's theorem. Thus the electric intensity at a point D near a uniformly charged sphere may be found by choosing our closed surface to be a sphere, concentric with the charged sphere and passing through D (Fig. 123). The area is $4\pi r^2$, and by symmetry the electric intensity is the same at every point of the sphere. Let it be E,

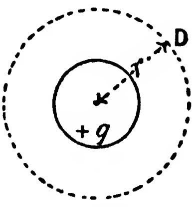


FIG. 123.

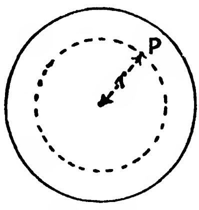


FIG. 124.

then total electric flux is $4\pi r^2 k\mathrm{E}$, and by Gauss's theorem it is also $4\pi q$;

$$\therefore \mathrm{E}=\frac{q}{kr^2}$$

that is, the strength of field at D is the same as though the charge q were all at the centre of the charged sphere.

The intensity inside a sphere throughout which there is a uniform density of charge of ρ units per unit volume can be found, for the sphere may be divided up into thin concentric

shells, and the effect of each shell which does not enclose P, the point at which the intensity is required, may be found by imagining the charge within it to be concentrated at the centre. The shells which enclose P do not add to the intensity at P (Fig. 124), as was seen on p. 113. The whole charge upon the shells which do not enclose P is equal to $\frac{4}{3}\pi r^3\rho$. And the intensity at P is therefore

$$\frac{4}{3}\cdot\frac{\pi r^3\rho}{kr^2}=\frac{4}{3}\cdot\frac{\pi r\rho}{k}.$$

The intensity is therefore greatest at the surface of the sphere, and falls off to zero at the centre.

Electric Intensity near Charged Cylinder.—In a similar manner the value of E due to an infinite cylinder which has a charge of q per centimetre length may be determined. We can see by symmetry that the field is everywhere radial and equal at equal distances from the axis. Hence, draw a coaxial cylinder through P (Fig. 125), and terminate it by two planes, unit distance apart, normal to the axis. E is parallel to these planes, and the electric flux over them is therefore zero. The area of the curved surface of the closed cylinder is $2\pi r$, and the total flux over the closed figure is therefore $2\pi rk$E. But the charge within it is q; therefore, by Gauss's theorem—

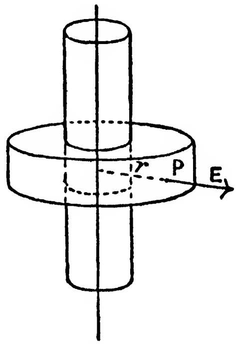


FIG. 125.

$$2\pi rkE=4\pi q$$

and,
$$E=\frac{2q}{kr}.$$

Since this is independent of the radius of the charged cylinder it holds also for a linear charge.

Intensity near Plane Sheet of Charge.—The value of E near an infinite plane sheet, having a charge of surface density σ_0 units per square centimetre, may be found by drawing a prism whose edges are normal to the plane, to cut the surface in unit area. If the plane is infinite in extent, we see by symmetry that the field is everywhere normal to the plane, and is of the same strength on each side of it. The charge within the prism is σ_0, and therefore the total electric flux over it is $4\pi\sigma_0$. The flux over the sides of the prism is zero, since they

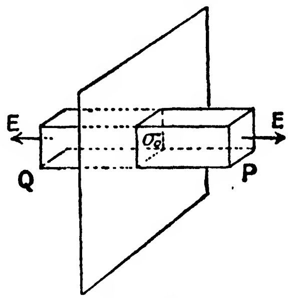


FIG. 126.

are everywhere parallel to E, and if the ends P and Q (Fig. 126) are planes parallel to the sheet, the area of each is unity, and they are both normal to E.

Then, total electric flux$=2k\mathrm{E}=4\pi\sigma_0$

$$\therefore \mathrm{E}=\frac{2\pi\sigma_0}{k}.$$

It will be noticed that this is independent of the distance from the sheet, and the electric intensity at any point near an infinite plane sheet is therefore $\frac{2\pi\sigma_0}{k}$. The charge is not situated upon a conductor, but is merely a sheet of charge with dielectric on both sides of it.

The electric intensity in the neighbourhood of a thin *charged plane conducting sheet* may be found from the last result by considering σ_0 to be the surface density of charge on both sides taken together. If these two happen to be equal, which in practice is unlikely to be the case, the surface density upon each side is $\frac{\sigma_0}{2}$. Calling this σ, we see that the intensity near a plane conductor on each surface of which the density of the charge is σ is equal to $4\pi\sigma$. It is, however, more satisfactory to establish this relation by drawing one of the closed ends Q of the prism, within the conductor, as in Fig. 127. Then the electric flux over Q is zero, since there is no electric intensity inside a conductor when the charges are at rest upon it. The whole flux $4\pi\sigma$ passes through P, and since this has unit area,

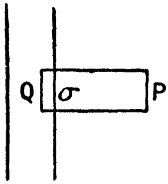


FIG. 127.

$$k\mathrm{E}=4\pi\sigma, \text{ or } \mathrm{E}=\frac{4\pi\sigma}{k}.$$

This is known as Coulomb's theorem.

Region inside a Conducting Surface.—We may also prove from Gauss's theorem that the *space within a closed equipotential surface is at uniform potential* when there is no electric charge within the surface. For let A (Fig. 128) be the equipotential surface; then if the space within A is not at the potential of A, a second equi-potential surface B can be drawn just inside A. The potential of B is either above or below that of A. If above, there is a field everywhere directed from B to A over the surfaces, and if below, this field is directed from A to B. In either case $\int k\mathrm{E}ds$ is equal to 4π times the charge within the surfaces, and since by hypothesis there is no such

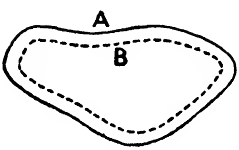


FIG. 128.

charge, $k\int E ds = 0$. But since in either case E has the same sign all over the surfaces, E must be zero at all points of the surface. That is, B is at the same potential as A. The same argument applies to the whole of the space within A, which is therefore at uniform potential, and there is no field within A.

It also follows that *there cannot be any charge on the inner surface of a hollow conductor* unless there is a charge in the hollow space within the conductor, and *if there be such a charge in the hollow space there will be an equal and opposite charge upon the inner surface of the closed conductor.* Consider B to be the inner surface of the conductor of which A is the outer surface (Fig. 128). Imagine a closed surface to be drawn between A and B and surrounding B. The intensity over this surface is everywhere zero, since the surface lies entirely in the conducting medium, and, by Gauss's theorem, the total charge within it is therefore zero. Hence, if there is any charge within B there must be an equal and opposite charge upon B, which establishes the second part of our proposition.

When there is no charge in the space within B there might still conceivably be equal and opposite charges upon different parts of B. Let a closed curve be drawn upon any part of B which may be supposed to have a charge upon it, and draw a closed surface to intersect B in this curve. The part of this closed surface within the conducting medium has no electric flux across it, for E is everywhere zero within the medium, and the part in the space within B has also no flux over it, as we have just seen, since E is zero within it. Hence the charge within our closed surface is zero, and that on the small area of B considered is also zero. That is, there is no charge at any point of B.

This proposition and the preceding one make general the problem proved on p. 113 for the sphere—that there is no field inside a conductor due to a charge outside it. It is also of importance in practice, for we see that a closed conductor constitutes a perfect electric screen for points inside it. Whatever the distribution of electric charge or intensity outside it, the conductor, since it is an equipotential surface, reduces the intensity inside it to zero.

Tubes of Induction and Lines of Force.—The important part played by the dielectric in electrical phenomena led Faraday to imagine tubes or lines of strain to exist in the medium situated between charged conductors, the positive charge upon one conductor and the negative charge upon the other being merely the ends of these tubes or lines. Maxwell gave these tubes a quantitative significance, and showed that the forces between the charges could be correctly represented by assuming the tubes to

be under a tension equal to $\frac{kE^2}{8\pi}$ in the direction of the tubes, and a pressure $\frac{kE^2}{8\pi}$ normal to them, and showed that such a system of tensions and pressures in a medium would be in equilibrium. Thus, the tubes tending to shorten, owing to this tension, would pull the opposite charges together, and the pressure of the tubes upon each other at right angles to their direction would push like charges apart (see Fig. 9).

Consider a small area S_1 (Fig. 129) drawn in an electric field with its plane at right angles to the direction of the field, and through its boundary let lines be drawn whose direction shall everywhere be that of the field. These lines enclose a tube, and if a second area S_2 be drawn anywhere at right angles to the field, it follows from Gauss's theorem, since there is no flux across the side of the tube, it being everywhere parallel to the field, that the flux over S_1 is equal to that over S_2.

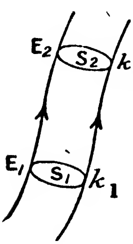


FIG. 129.

Thus $k_1E_1S_1=k_2E_2S_2$, provided that there is no charge in the element of space situated between S_1 and S_2, and so the electric flux over any section of the tube is a constant quantity. Calling the flux per unit area ϕ, we have $kE=\phi$, and $\phi_1S_1=\phi_2S_2$.

If the tube be traced back to the charged conductor upon which it arises, we can, by applying Gauss's theorem, as in proving Coulomb's theorem (p. 126), show that $4\pi(\sigma S)=\phi S$, where σ is the surface density of the charge upon the conductor, and σS is therefore the charge upon which the tube arises. If this is a unit charge, the tube is said to be a unit or Faraday tube, and we see that the number of Faraday tubes is numerically equal to the charge upon which they arise, each unit of charge giving rise to one tube. Thus with a surface density of charge σ there are σ Faraday tubes arising upon each square centimetre of surface. Calling D the number of Faraday tubes per square centimetre, we see from the last equation that—

$$4\pi D=\phi=kE$$

$$\therefore D=\frac{kE}{4\pi}$$

D is called by Maxwell the electric displacement in the medium, by which he means, the amount of electricity which is caused to cross each unit of area of the dielectric on account of the electric intensity at that point. Thus, if there is an electric intensity (which Maxwell calls an electromotive force) acting in a conductor, the charge continually moves on account of it; but in a

dielectric this motion is not indefinite; the displacement reaches a limiting amount which is proportional to the force producing it. Thus the displacement and intensity are related to each other in the electrical case, like the strain and stress in the case of elasticity. Since a tube of induction starts upon a positive and ends upon a negative charge, the positive charge may be looked upon as a displacement in one direction at one end of the tube, and a negative charge as an opposite displacement at the other end of the tube, and throughout the tube the displacement is continuous, but its value changes with the area of cross-section of the tube. Some writers use the name of polarisation for this quantity, as it is of a more general character than "displacement," but this name is used here for the electric moment of unit volume of a dielectric.

From the relation $\phi_1 S_1 = \phi_2 S_2$, or $D_1 S_1 = D_2 S_2$, we see that for a given tube of induction, the number of Faraday tubes per unit area is inversely as the cross-section of the tube; also, since $E = \frac{4\pi D}{k}$, we see that so long as k is constant at all parts of the medium E is proportional to D, *i.e.* to the number of Faraday tubes per unit area. If, then, instead of representing the field by tubes, we draw a line down the middle of each tube, the number of such lines per square centimetre is equal to D. It is more usual, however, to draw 4π lines for every Faraday tube, and the number per unit area is then equal to ϕ. These are called lines of induction. If, further, $k=1$, the number of lines per square centimetre is equal to the strength of field E. Such lines are called lines of force. This conception of lines of force and lines of induction is a very useful one for the graphical representation of fields of force, and is used very frequently in the case of magnetic as well as electrical problems.

Energy in Medium.—From the analogy with problems in elasticity, we should expect that in a dielectric there is an amount of energy per unit volume corresponding to the quantity $\frac{1}{2}$(stress $\times$ strain). Consider an element of a tube of induction whose length dl is so small that the electric intensity, and therefore the area of cross-section, may be considered to be constant. SS, being at right angles to E, are equipotential surfaces (Fig. 130), and therefore if we arrange two conducting surfaces, one to coincide with each, and each having the appropriate potential, we shall not alter the problem with respect to the space between them (see p. 122). Upon one of these there will be a surface density of charge $+\sigma$ and upon the other $-\sigma$, and $E = \frac{4\pi\sigma}{k}$, the charges at

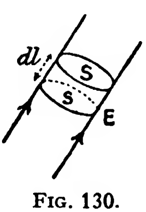


FIG. 130.

the ends of the tubes being $+\sigma S$ and $-\sigma S$ respectively. The force upon unit charge is E, and hence work done in carrying unit charge from one end to the other is $E\,dl$, which is therefore the difference of potential V between them (see p. 119). But energy$=\frac{1}{2}QV$, and Q in this case is σS;

$$\therefore \text{ energy}=\tfrac{1}{2}\,.\,\sigma S\,.\,E\,.\,dl$$
$$=\frac{1}{2}\,.\,\frac{kE^2}{4\pi}\,.\,S\,dl\,.=\frac{kE^2}{8\pi}S\,.\,dl.$$

$S\,dl$ is the volume of the element, and therefore the energy per unit volume is $\dfrac{kE^2}{8\pi}$,

or since, $$E=\frac{4\pi D}{k},$$

$$\text{energy per unit volume}=\frac{2\pi D^2}{k}=\tfrac{1}{2}ED.$$

Thus, if E is of the nature of a stress, D is the corresponding strain.

It may be noted that in calculating the energy, the charge at one end only of the tube is used. This is in accordance with our procedure in calculating the energy $\frac{1}{2}QV$ for a given charge Q (p. 122). For Q is the charge placed on the given conductor and although there must be an equal and opposite charge at the other end of the tubes of induction arising upon Q, we did not take this opposite charge into the term Q used in calculating the energy.

Force on Surface of Charged Conductor.—The expression for the energy in the medium might have been obtained by first calculating the outward force per unit area acting normally upon a charged conducting surface, and then imagining the surface displaced through a small distance. Let p be a point very close to a conductor upon which the surface density of charge is σ (Fig. 131). The electric intensity E may be considered to consist of two parts, f due to the charge situated in the neighbourhood of p, and f' due to all other charges. Then $f+f'=E$. At the point p', inside the conductor and indefinitely close to p, f' is the same as at p, but f is reversed in sign since p' is situated on the opposite side of the neighbouring charges to p, which charges are of course on the surface of the conductor. The resultant intensity is therefore $f'-f$. But this is zero, since p' is inside the conductor,

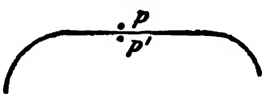


FIG. 131.

$$\therefore f=f'=\frac{E}{2}, \text{ a result originally due to Laplace.}$$

Now, the charge σ upon unit surface is situated in the field $f'=\frac{E}{2}$, and the force on it is therefore $\frac{\sigma E}{2}$.

$$\text{But, } E=\frac{4\pi\sigma}{k},$$

$$\therefore \text{ force per unit area of surface}=\frac{2\pi\sigma^2}{k}=\frac{2\pi D^2}{k}=\frac{kE^2}{8\pi}.$$

This result might have been deduced from a consideration of the polarisation of the medium between two parallel plates (Fig. 117, p. 117) together with the field that would exist in empty space. The field on the upper side of the plate A is normal to the plane and of value $E=\frac{4\pi\sigma}{1+4\pi\epsilon}$. This may be considered to consist of two parts $\frac{E}{2}$ due to A and the same due to B. The charge σ per unit area of A is then situated in the field $\frac{2\pi\sigma}{1+4\pi\epsilon}$ due to B and therefore experiences the outwardly directed force $\frac{2\pi\sigma^2}{1+4\pi\epsilon}=\frac{2\pi\sigma^2}{k}=\frac{kE^2}{8\pi}$. The distance between the plates is immaterial since the electric field is uniform, a condition ensured by the presence of B.

If, then, the surface be displaced in the direction of its normal, through distance dl, work done per unit area of surface$=\frac{kE^2}{8\pi}dl$; but the volume swept out by unit area is dl, therefore work done in producing unit volume of electric field is $\frac{kE^2}{8\pi}$, which is therefore the energy associated with unit volume of the dielectric.

Stresses in Tubes of Induction.—On the assumption that electrostatic phenomena are due to stresses in the medium, we should expect, from the fact that there is a pull upon a charged conductor equal to a tension of $\frac{2\pi D^2}{k}$, that this pull is due to tension in the tube itself. It follows that if the tubes are in a state of tension they must also exert a lateral pressure upon neighbouring tubes, since if this were not the case, the tubes passing from a small positive charge to a similar negative one, would shrink until they became straight lines joining the charges, and the rest of the medium would be entirely free from them. As this is not the case, we must assume that they exert a lateral pressure upon each other, and we will now find the value of this

pressure which is necessary to produce equilibrium with the tension in the tubes.

Consider a small section of a tube of induction, the sides AE, BF, CG and DH (Fig. 132) being parallel to the field, and the ends ABCD and EFGH equipotential surfaces. The forces f_1 and f_2 on the faces ABCD and EFGH are due to the tensions $\frac{2\pi D_1^2}{k}$ and $\frac{2\pi D_2^2}{k}$ at the respective faces.

Then,
$$f_1 = \frac{2\pi D_1^2}{k} \cdot a_1$$
and
$$f_2 = \frac{2\pi D_2^2}{k} \cdot a_2,$$

where a_1 and a_2 are the areas of the faces, and D_1 and D_2 the corresponding displacements,

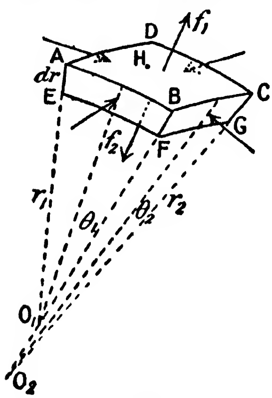


FIG. 132.

$$\therefore f_1 - f_2 = \frac{2\pi}{k}(D_1^2 a_1 - D_2^2 a_2)$$
$$= \frac{2\pi}{k}(D_1 D_2 a_2 - D_1 D_2 a_1)$$

since $D_1 a_1 = D_2 a_2$ (p. 129),
$$\therefore f_1 - f_2 = \frac{2\pi D_1 D_2}{k}(a_2 - a_1).$$

Since $a_1 > a_2$, it follows that $f_2 > f_1$, and there is a resultant force acting in the direction of f_2. Also, if the thickness of the slice is small, we may write D^2 in place of $D_1 D_2$.

$\therefore$ resultant force in the direction of f_2 is $\frac{2\pi D^2}{k}(a_1 - a_2)$.

If O_1 is the centre of curvature of AB and EF, r_1 the radius of curvature of the sides EF and GH, and θ_1 the semi-angle subtended at the centre of curvature by EF,
$$EF = 2r_1\theta_1, \text{ and, } AB = 2(r_1 + dr)\theta_1,$$
In a similar manner, $FG = 2r_2\theta_2$, and $BC = 2(r_2 + dr)\theta_2$, so that $a_2 = 4r_1r_2\theta_1\theta_2$, and $a_1 = 4(r_1 + dr)(r_2 + dr)\theta_1\theta_2$.

Neglecting the small quantity $(dr)^2$, we have—
$$a_1 - a_2 = 4(r_1 + r_2)\theta_1\theta_2 \,.\, dr,$$
$$\text{and, } f_2 - f_1 = \frac{2\pi D^2}{k} \cdot 4(r_1 + r_2)\theta_1\theta_2 \,.\, dr.$$

For the section to be in equilibrium, the pressures over the sides must produce a resultant force in the direction f_2, equal and opposite to the above.

Again, if p be the lateral pressure; force over side ABFE $=p \cdot 2r_1\theta_1 \cdot dr$, and this is inclined at angle $\left(\frac{\pi}{2}-\theta_2\right)$ to f_2. Hence component parallel to f_2 is $2 \cdot pr_1\theta_1 \cdot dr \cdot \sin\theta_2$. But if the element is small θ_2 may be written for $\sin\theta_2$, so that—

force parallel to f_2, for side ABFE, $=2pr_1\theta_1\theta_2 dr$,
and for the two opposite sides taken together $=4pr_1\theta_1\theta_2 dr$.

In an exactly similar manner we see that the component due to the pair of sides BCGF and ADHE is $4pr_2\theta_1\theta_2 dr$,

$\therefore$ resultant force parallel to $f_2=p \cdot 4(r_1+r_2)\theta_1\theta_2 dr$.

Comparing this with the value of f_2-f_1, we see that for equilibrium—

$$p=\frac{2\pi D^2}{k}.$$

When the two ends of an element of a tube are parallel to each other, we have just seen that the stresses over the ends are in equilibrium with the pressures over the sides; but in an electric field the tubes are not in general straight, and where they are curved, the tensions over the two ends of an element of the tube have a resultant at right angles to the tube, and it is necessary for us to see whether the tube is still in equilibrium under the equal longitudinal tension and lateral pressure $\frac{2\pi D^2}{k}$. If the tube is curved, let the plane of the diagram (Fig. 133) be taken through the direction of curvature.

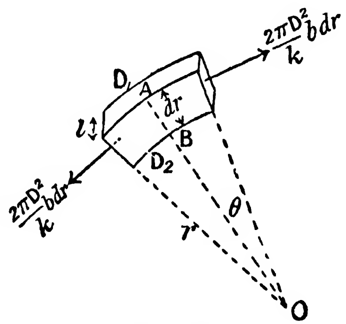


FIG. 133.

Let the section of the tube considered be short enough for us to take D as constant over its length, then the tension $\frac{2\pi D^2}{k}$ at each end will give rise to forces $\frac{2\pi D^2}{k} \cdot bdr$, and since these are equally inclined to the median line BO, and θ is small—

$$\text{Resultant force along BO}=2\frac{2\pi D^2}{k}b \cdot dr \cdot \theta.$$

D is the mean displacement over one end, but it is different at the outer and inner sides A and B; for, the ends being equipotential surfaces, the p.d. as measured along A from one face

to the other is equal to that as measured along B; that is—

$$E_1(r+dr)\theta=E_2r\theta.$$

But, $$E_1=\frac{4\pi D_1}{k}, \text{ and, } E_2=\frac{4\pi D_2}{k},$$

$$\therefore D_1 . (r+dr)=D_2 . r.$$

Thus p_1, the lateral pressure over A, is $\frac{2\pi D_1^2}{k}$, and the resultant force over the outer face A is $\frac{2\pi D_1^2}{k}2(r+dr)\theta . b$, and that over B is $\frac{2\pi D_2^2}{k}2r . \theta . b$.

Since these both act in the line BO—

$$\text{Resultant force on element}=\frac{4\pi}{k} . \theta b\{D_1^2(r+dr)-D_2^2r\}$$

$$=\frac{4\pi}{k}\theta b\{D_1D_2r-D_1D_2(r+dr)\}$$

$$=-\frac{4\pi D_1D_2}{k} . \theta b . dr.$$

This is directed outwards, and writing D^2 instead of D_1D_2, we see that this is equal and opposite to the resultant of the forces over the ends due to the tension, and the element is in equilibrium. Since the result is obtained on the assumption that $p_1=\frac{2\pi D_1^2}{k}$, and $p_2=\frac{2\pi D_2^2}{k}$, we see that our assumption is justified.

Motion of Tubes, and Electric Current.—The phenomenon of the electric current may be explained in terms of these tubes of induction. Electric charges are the ends of the tubes, and these are free to move upon the conductor, and will therefore slide along it until the tubes in the surrounding dielectric are in equilibrium. Thus, if the conducting plates A and B (Fig. 134) are oppositely charged, the ends of the tubes will slide along A and B until equilibrium is established. If, then, the plates are connected by a conductor CD, the opposite ends of the tubes nearest to CD can approach each other, and owing to their tension these tubes will contract until they vanish. This removes the lateral pressure to the left of the tube EF, and hence the pressure upon the right will push this tube towards CD, and it will in turn vanish. The

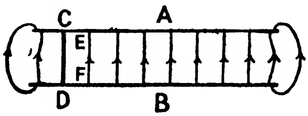


FIG. 134.

process will go on until all the tubes have disappeared. The motion of the positive ends along B and the negative ends along A constitutes the current.

The magnetic field in the neighbourhood of an electric current has been interpreted by Sir J. J. Thomson in terms of the lateral motion of the Faraday tubes (Chap. XIII).

Distribution of Charge upon a Conductor.—The fact that the surface of a conductor must be one of equipotential aids us in determining the way in which a charge is distributed upon it. In the case of a symmetrical surface, such as a sphere or an infinite plane, the problem presents no difficulty, the lateral pressures due to the tubes of induction ensuring a uniform distribution of charge. When this symmetry is departed from, the problem of finding exactly the distribution of charge and field presents great difficulties, but general reasoning will show us that on any given conductor in an open space the charge is distributed so that the surface density is highest on parts of greatest convexity, becoming infinite at an actual point.

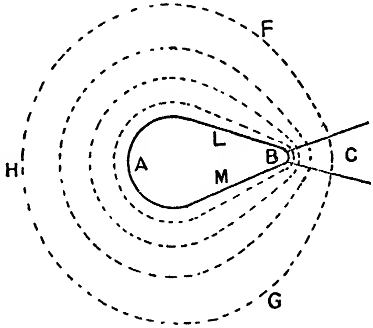


FIG. 135.

Taking a conductor of the form AB (Fig. 135) and drawing the equipotential surfaces in its neighbourhood, we see that near to it they follow its outline very closely, since the surface of the conductor itself is an equipotential surface. Owing to the great curvature of the equipotential surfaces in the neighbourhood of a point such as B, a tube of induction such as BC has a very great divergence, that is, its cross-section varies rapidly as we pass from C and B. Now, for any given tube the product of electric displacement and area of section, that is, DS, is constant (p. 129), and therefore, as S becomes very small on approaching B, D necessarily becomes very great.

At a distance from the conductor, the equipotential surfaces, such as FG, are approximately spherical, and each unit tube of induction has here the same cross-section. The convergence towards B being greater than that towards A, the cross-section of a unit tube is less at B than at A, and the electric intensity at B is therefore greater than at A.

In the event of B being a point, S becomes zero and D infinite. Remembering that $E = \frac{4\pi D}{k} = \frac{4\pi\sigma}{k}$, we see that the electric intensity in the neighbourhood of a point, and the surface density

of charge on the point, are both infinite. But long before this condition is reached, the insulation of the air or other dielectric surrounding the conductor breaks down, and the charge passes from the point.

It has long been known that fine points facilitate the discharge of a conductor, and produce what is called an electric wind. This discharge from fine points has been used for many purposes, as in collecting the charge from the sectors of an electrical machine (p. 111).

A further examination of the equipotential surfaces of Fig. 135 shows us that we can easily obtain an idea of the surface density at all points of the conducting surface, and of the electric intensity of the field, for we cross the same number of equipotential surfaces in going from the conductor to the surface FG by whatever path we go, and therefore the longer the path the less close together are the surfaces. This means a less potential gradient and a weaker field. From B to C is the shortest path, and here we find the strongest field.

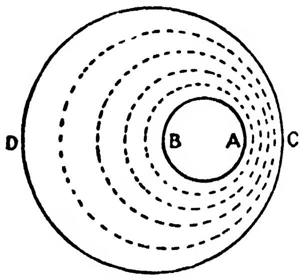


FIG. 136.

A similar method enables us to follow the effect of want of symmetry of the surrounding conductors upon a body itself symmetrical. For, taking a charged sphere AB inside a conducting sphere; both spheres are equipotential surfaces, and therefore in passing from A to C (Fig. 136) we cross as many equipotential surfaces as in going from B to D. Hence the field between A and C is stronger than that between B and D, and in the approximate ratio of the distances BD : AC. Remembering that $E=\frac{4\pi\sigma}{k}$, we see that the surface density of charge upon A is greater than that upon B.

Force on Uncharged Body.—The force on an uncharged body situated in an electric field may be determined in direction by a simple consideration of the energy of the field. In a uniform field of intensity of induction ϕ, the energy per unit volume is $\frac{kE^2}{8\pi}=\frac{\phi^2}{8\pi k}$, since $\phi=kE$, so that for a body of dielectric constant k situated in the field, $\frac{\phi^2}{8\pi k}$ is the energy per unit volume of the body; whereas if the space occupied by the body had been occupied by air the energy per unit volume would have been $\frac{\phi^2}{8\pi}$.

Since k is greater than unity, the energy of the field when occupied by a body is less than when the space is occupied by air, and for such a body situated in air, the energy of the whole system of air and body is less when the body is present than when it is not. If the field be uniform, the energy is the same wherever the body may be situated, and there is consequently no tendency for the body to move from one place to another. If, however, the field is not uniform, the energy of the whole system is less when the body occupies a position where ϕ is great than in one where ϕ is not so great. Now, it is a general principle in dynamics that a system will always tend to that configuration for which the total potential energy is least, so that in our case the body will experience a force urging it from points of weaker to points of stronger field. Owing to the difficulty of determining the distribution of ϕ in the case of a body situated in a field which is not uniform, we cannot, as a rule, employ the above reasoning to calculate the actual force on the body, but the general principle holds that the force acts towards the place of greatest field.

In the case of a conducting body, the induction ϕ inside it is zero. The energy within the body is therefore zero, and from the above principle of least potential energy, the conductor will experience a force urging it from weaker to stronger parts of the field.

The above reasoning explains why a charged body will attract an uncharged one, as in the case of the rubbed amber or ebonite attracting light bodies, such as the pith ball.

Boundary Conditions.—At the surface of separation of two different dielectrics, certain conditions must hold, which conditions may be obtained in quite a simple manner. For convenience we shall first study the condition applying to a field whose direction is parallel to the surface of separation, and then consider the case of a field normal to this surface.

(i) *Field Parallel to Surface of Separation.*—Let k_1 and k_2 be the dielectric constants of the two media, and E_1 and E_2 the fields in each. Draw two equipotential surfaces A and B (Fig. 137) through the surface of separation and an infinitesimal distance dl apart. Then A and B must be parallel, since equipotential surfaces are always perpendicular to the field, and, further, the potential difference $V_A - V_B$ is for the first medium $E_1 dl$, and for the second $E_2 dl$.

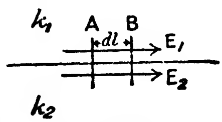


FIG. 137.

Since these are equal, $E_1 = E_2$, so that our first boundary condition is, that ***the tangential components of the electric intensity are the same on the two sides of the surface of separation.***

(ii) *Field normal to Surface of Separation.*—In this case, we take a small closed surface with ends parallel to the boundary and sides normal to it. Then, if ϕ_1 and ϕ_2 are the normal inductions taken positive in the direction from medium 2 to medium 1 (Fig. 138), the total electric flux over a_1 is $\phi_1 a_1$, and over a_2 is $\phi_2 a_2$, and if a charge q be situated on the boundary we have, from Gauss's theorem—

$$\phi_1 a_1 - \phi_2 a_2 = 4\pi q$$

or since $a_1 = a_2$—

$$\phi_1 - \phi_2 = 4\pi\sigma, \text{ because } \sigma = \frac{q}{a}.$$

In the particular case when there is no charge upon the surface of separation—

$$\sigma = 0, \text{ and, } \phi_1 = \phi_2$$

thus the second condition is that *the normal component of the induction is the same in both media.*

Since $\phi = kE$, condition (ii) may be written $k_1 E_1 = k_2 E_2$.

These two conditions are of great importance in the study of

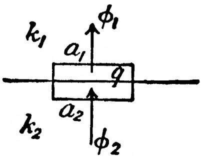


FIG. 138.

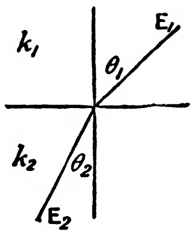


FIG. 139.

the problem of the reflection of electromagnetic waves by a surface of discontinuity.

The two boundary conditions enable us to find the change in direction of the field as we pass from one medium to another. For let θ_1 be the angle between field and normal (Fig. 139) in the first medium, and θ_2 that in the second.

The first boundary condition gives us

$$E_1 \sin\theta_1 = E_2 \sin\theta_2,$$

and the second, $k_1 E_1 \cos\theta_1 = k_2 E_2 \cos\theta_2$.

Therefore dividing one equation by the other—

$$\frac{\tan\theta_1}{\tan\theta_2} = \frac{k_1}{k_2}.$$

In the equation $\phi_1 - \phi_2 = 4\pi\sigma$, or $k_1 E_1 - k_2 E_2 = 4\pi\sigma$, σ is a charge which may be placed upon the surface or removed from it, and does not owe its existence to the discontinuity at the

surface of separation. When this is zero, $k_1E_1=k_2E_2$. Maxwell considered the difference in field E_1-E_2 on the two sides of the medium to be due to a fictitious charge upon the boundary, which will, of course, disappear when the inductions ϕ_1 and ϕ_2 disappear. Thus, if the dielectric constants upon both sides of the surface become equal to unity, the intensities will remain unchanged, provided that this fictitious surface charge σ' remains upon the surface, σ' being obtained by putting $k_1=k_2=1$ in the above equation.

Then,
$$E_1-E_2=4\pi\sigma'$$

Remembering that $k_1E_1=k_2E_2$ when there is no other than the fictitious charge on the surface, we see that—

$$4\pi\sigma'=\frac{k_2}{k_1}E_2-E_2=\frac{k_2-k_1}{k_1}E_2$$
$$=E_1-\frac{k_1}{k_2}E_1=\frac{k_2-k_1}{k_2}E_1.$$

If k_2 becomes equal to k_1, which is not unity, we still obtain σ' as before. But in this case—

$$4\pi\sigma'=k_1(E_1-E_2).$$

Thus for any given problem the intensities on the two sides of the boundary will be unchanged if we change the original dielectric constants k_1 and k_2 to unity and add a surface density of charge

$$\sigma'=\frac{1}{4\pi}\cdot\frac{k_2-k_1}{k_2}E_1=\frac{1}{4\pi}\cdot\frac{k_2-k_1}{k_1}E_2.$$

Uncharged Sphere in Electric Field.—By the aid of this idea of a fictitious surface density we can solve the problem of the distribution of intensity in the case of a sphere situated in a medium whose dielectric constant differs from that of the sphere.

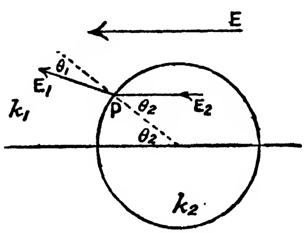


FIG. 140.

Let the dielectric sphere of constant k_2 be situated in a medium of dielectric constant k_1, and let the field be uniform before the introduction of the sphere, the intensity being everywhere E. When the dielectric constant of the sphere changes to k_2, we have for the surface condition for the normal component of the induction at the point P, Fig. 140, the relation $k_1E_1\cos\theta_1=k_2E_2\cos\theta_2$, where E_1 and E_2 are fields just outside and inside the sphere at P. We may now produce exactly the same fields normal to the surface of the

sphere, if we make the dielectric constant everywhere k_1 and introduce a fictitious surface density σ' given by

$$k_1E_1 \cos\theta_1 - k_1E_2 \cos\theta_2 = 4\pi\sigma' \text{ (p. 139)},$$

or,
$$k_2E_2 \cos\theta_2 - k_1E_2 \cos\theta_2 = 4\pi\sigma'$$

$$\sigma' = \frac{E_2 \cos\theta_2}{4\pi}(k_2 - k_1).$$

The dielectric constant being now everywhere k_1, we have everywhere the original field E, together with that due to the fictitious surface density σ'.

A distribution of surface density which may be made to satisfy the conditions of the problem was suggested by Poisson. Let the sphere be considered to have two volume densities of charge, $+\rho$ and $-\rho$, which coincide when there is no external field; but the sphere of positive charge is displaced relatively to the sphere of negative charge, in the direction of the field, by the amount OO′=AA′=BB′ (Fig. 141), owing to the field E. MANA′ is then a layer of positive charge, and MBNB′ a similar layer of negative charge, and throughout the rest of the sphere the

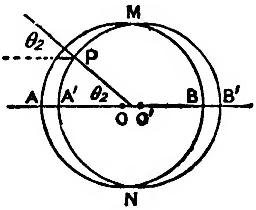


FIG. 141.

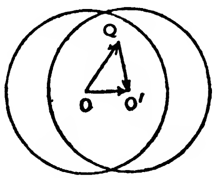


FIG. 142.

charges neutralize each other. When this displacement is very small, the surface density at a point P is represented by the length of a radius intercepted between the spheres, which length is OO′ $\cos\theta_2$.

Then the surface density at P is ρ . OO′ $\cos\theta_2$.

To find the field due to this distribution of charge, all that is necessary is to find that due to the two spheres. At a point Q (Fig. 142) the intensity due to the sphere, whose centre is O, which has volume density of charge $+\rho$, is equal to $\frac{1}{k_1}\frac{4}{3}\pi$. OQ . ρ (see p. 125), and may be represented by the vector OQ; that due to the other sphere is $\frac{1}{k_1}\frac{4}{3}\pi$. O′Q . ρ, and is represented by QO′. The resultant intensity may therefore, by the triangle of forces, be represented by the vector OO′, and the intensity is $\frac{1}{k_1}\frac{4}{3}\pi$. OO′ρ; and this is parallel to OO′ and is independent of the position

of Q. Thus the field inside the sphere is uniform and parallel to the original field. By finding the value of the intensity due to this distribution everywhere, and combining it with E, we get the resultant field everywhere.

Within the sphere, the field due to the charges is opposite in direction to the original field, so that the resultant field E_2 is the difference between these, is everywhere parallel to E, and is constant. For the boundary condition at the surface of the sphere to be satisfied,

$$\sigma' = \frac{E_2 \cos \theta_2}{4\pi}(k_2 - k_1)$$

and since the surface density due to the volume distributions is $\rho \,.\, OO' \,.\, \cos \theta_2$, we have

$$\rho \,.\, OO' = \frac{E_2(k_2 - k_1)}{4\pi}.$$

And again, since, $E - E_2 = \frac{1}{k_1} \,.\, \frac{4}{3}\pi \,.\, OO' \,.\, \rho,$

$$\frac{E_2}{4\pi}(k_2 - k_1) = \frac{3}{4\pi} \,.\, k_1(E - E_2),$$

$$E_2(k_2 - k_1) + 3E_2 k_1 = 3k_1 E,$$

$$E_2 = \frac{3k_1}{k_2 + 2k_1} E.$$

The field outside the sphere may be found by combining the uniform field E with that due to the charges $+\frac{4}{3}\pi a^3 \rho$ situated at O, and $-\frac{4}{3}\pi a^3 \rho$ situated at O′, remembering that the dielectric has a constant k_1.

In the case of a conducting sphere situated in air, $k_2 = \infty$ and $k_1 = 1$.

$$\therefore\ E_2 = \frac{E}{\infty} = 0,$$

which is in accordance with fact, since the intensity inside a conductor is zero. Also when $\theta_2 = 0$,—

$$\sigma' = OO' \,.\, \rho = \frac{3}{4} \,.\, \frac{k_1}{\pi}(E - E_2)$$

and since $E_2 = 0$, and $k_1 = 1$—

$$\sigma' = \frac{3}{4\pi} E, \text{ at the points A and B (Fig. 141).}$$

At any other point of the sphere $\sigma' = \frac{3}{4\pi} E \cos \theta_2$.

It may be shown that the field inside an ellipsoid with one axis parallel to E is also uniform, and the case of the sphere deduced from that of the ellipsoid by making the axes equal. The problem

presents mathematical difficulties which preclude the discussion of it here, but the student may find it in "Absolute Measurements in Electricity and Magnetism," by A. Gray.

From analogy with the case of a small magnet (p. 5) we can

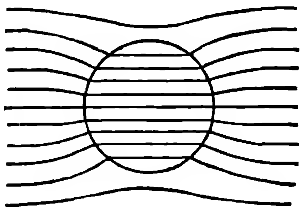


FIG. 143.

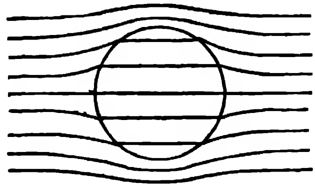


FIG. 144.

see that the field outside the sphere is equal to that due to a doublet of moment

$$\tfrac{4}{3}\pi a^3\rho \,.\, \mathrm{OO}', \text{ or, } a^3k_1(\mathrm{E}-\mathrm{E}_2)=a^3k_1\left(1-\frac{3k_1}{k_2+2k_1}\right)\mathrm{E}$$

$$=a^3\frac{k_1(k_2-k_1)}{k_2+2k_1}$$

Fig. 143 represents the direction of the resultant field or induction for a sphere, when $k_2>k_1$. In Fig. 144, $k_2<k_1$.

Electrical Images.—*Conducting Plane.*—The distribution of charge upon conductors may, in several cases, be found most simply by the method of electrical images, due to Lord Kelvin. Let us find the distribution of charge over a plane conducting surface, due to the presence of a positive charge $+q$ situated at P (Fig. 145), when the plane is maintained at zero potential. At P′, a point on the opposite side of the plane to P, so that PL=P′L, and PLP′ is normal to the plane, place a charge $-q$. This charge is said to be the electrical image of P, the name being suggested by the optical analogy.

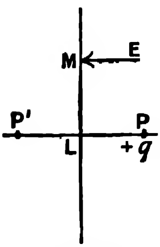


FIG. 145.

We must show that this analogy may be pushed further; in fact, for all points on the P side of the plane the effect of the plane is exactly the same as that produced by the image, the plane being removed.

In the first place, the potential at M is

$$+\frac{q}{\mathrm{PM}}-\frac{q}{\mathrm{P'M}}=0,$$

so that if the conductor were removed and the charge $-q$ placed at P′ instead, every point of the plane LM would still be at zero potential.

Again, the electric intensity E, at the point M is given by

$$E=\frac{q}{PM^2}\cdot\frac{PP'}{PM} \text{ (compare with field due to magnet, p. 4)}$$
$$=\frac{2q\,.\,PL}{PM^3}$$

But if $-\sigma$ is the surface density of charge at M, we have from Coulomb's theorem (p. 126)

$$E=-4\pi\sigma$$

and if these two are to be the same,

$$-4\pi\sigma=\frac{2q\,.\,PL}{PM^3}$$
$$\sigma=-\frac{q\,.\,PL}{2\pi PM^3}$$

which determines the value of σ at all points on the plane. The two distributions $+q$ and $-q$ on the one hand, and $+q$ and $-\sigma$ on the other, both make LM a surface of zero potential, and both produce the same intensity at points immediately in contact with the plane, and since the component of the intensity due to P at any point is the same in both cases, it follows that the intensity near the plane due to $-q$ at P′ is identical with that due to $-\sigma$ on the plane. Moreover, it follows that if $+q$ and $-q$ on the one hand, and $+q$ and the distribution $-\sigma$ on the other, each make the potential of the plane zero, the intensity at every point on the P side of LM due to the two distributions must be the same. For if possible let the intensity due to the former arrangement be E_1 and that due to the latter E_2. In the second case reverse all the charges, E_2 becoming, of course, $-E_2$, with charge $-q$ at P and distribution $+\sigma$ upon the plane. If we now superimpose the second distribution on the first, the charge is everywhere zero, and the field at the given point is E_1-E_2. But the space to the right is surrounded by an equipotential surface made up of the plane LM and the rest of the enclosure at infinity, and as there is no charge within it, it follows from the theorem on p. 127 that the intensity within it must be everywhere zero, that is, $E_1-E_2=0$ or $E_1=E_2$, and thus the intensity due to the two distributions is everywhere the same.

Further, we can show that if the given distribution $-\sigma$, which produces a field everywhere on the P side of LM, produces an intensity equal to that due to $-q$ at P′; no other surface distribution can do so. For if possible let another $-\sigma'$, whose value is not everywhere the same as $-\sigma$, give the same intensity as $-\sigma$. Let this latter charge $-\sigma'$ be reversed and superimposed upon $-\sigma$. The intensities now cancel out everywhere, but the

charges do not. Thus at a given point on the plane, the resulting surface density is $\sigma'-\sigma$, and on drawing a closed surface to enclose this, the normal induction over it is $4\pi(\sigma'-\sigma)$, but P does not lie within the surface, and therefore by Gauss's theorem $+q$ cannot contribute to the total electric flux over it, and the field due to σ and σ' is everywhere zero; hence the total electric flux over this closed surface is zero, and $\sigma=\sigma'$. That is, there is only one distribution of σ that can satisfy the problem, and since one has been found, it must therefore be the only one possible.

Thus the charge $+q$ situated at P produces a surface density $-\frac{q\text{PL}}{2\pi\text{PM}^3}$ at the point M of the conducting plane. There is evidently an attraction between $+q$ and the negatively charged plane, the value of which can be found by replacing the charge on the plane by the electrical image of P.

$$\text{Force}=-\frac{q^2}{(2\text{PL})^2}=-\frac{q^2}{4(\text{PL})^2}.$$

Electrical Images.—*Conducting Sphere.*—The only other case of an electrical image which we shall consider is the image of a charge $+q$, produced by a spherical conducting surface at zero potential. Let the point P′ within the sphere of radius r (Fig. 146) be found such that $\text{OP}\,.\,\text{OP}'=r^2$.

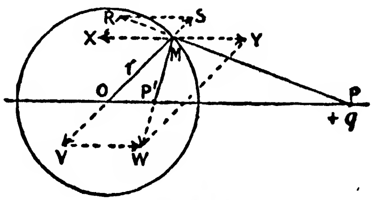


FIG. 146.

Then, $$\frac{\text{OP}}{\text{OM}}=\frac{\text{OM}}{\text{OP}'}$$

Thus the triangles OPM and OMP′ are similar, and

$$\frac{\text{PM}}{\text{P}'\text{M}}=\frac{\text{OP}}{\text{OM}}=\text{constant}.$$

If the charge $-\frac{\text{P}'\text{M}}{\text{PM}}q$ be placed at P′,—

$$\text{potential at M due to } +q \text{ at P}=+\frac{q}{\text{PM}},$$

$$\text{and potential at M due to} -\frac{\text{P}'\text{M}}{\text{PM}}q \text{ at P}'=-q.\frac{\text{P}'\text{M}}{\text{PM}}\cdot\frac{1}{\text{P}'\text{M}}$$

$$=-\frac{q}{\text{PM}}.$$

Therefore the two charges together produce zero potential at every point on the sphere. As in the previous case $-q \cdot \frac{\text{P'M}}{\text{PM}}$ placed at P′ is the electrical image of $+q$ at P, since it reduces the potential at any point of the surface, such as M, to zero.

Taking OP$=d$, we have—

$$-q \cdot \frac{\text{P'M}}{\text{PM}} = -q\frac{\text{OM}}{\text{OP}} = -\frac{rq}{d}.$$

The force between the charge and the sphere is found as in the last case, by taking the image in place of the charge on the sphere—

$$\text{F} = -\frac{rq}{d} \cdot q \cdot \frac{1}{(\text{PP'})^2}$$

$$= -\frac{rq^2}{d(d-\text{OP'})^2}.$$

$$\text{But OP'} = \frac{r^2}{d},$$

$$\therefore \text{F} = -\frac{rq^2}{d\left(d-\frac{r^2}{d}\right)^2} = -\frac{q^2rd}{(d^2-r^2)^2}.$$

To find the surface density of charge $-\sigma$ at the point M, resolve the intensity $\frac{q}{\text{PM}^2}$ along OM, and parallel to OP. From Fig. 146, SMR is the triangle of forces, and is similar to OMP, so that the component MS along the radius is equal to

$$\frac{q}{\text{PM}^2} \cdot \frac{\text{MS}}{\text{MR}} = \frac{q}{\text{PM}^2} \cdot \frac{\text{OM}}{\text{PM}} = \frac{qr}{\text{PM}^3}.$$

Also the triangles MVW and MOP′ are similar, and the component of MW, or $q\frac{\text{P'M}}{\text{PM}} \cdot \frac{1}{\text{P'M}^2}$, in the direction of the radius OM is

$$q \cdot \frac{\text{P'M}}{\text{PM}} \cdot \frac{1}{\text{P'M}^2} \cdot \frac{\text{MV}}{\text{MW}} = \frac{q}{\text{PM} \, . \, \text{P'M}} \cdot \frac{\text{OM}}{\text{P'M}} = \frac{qr}{\text{PM} \, . \, \text{P'M}^2} = \frac{qd^2}{r \, . \, \text{PM}^3}$$

Similarly the components MX and MY parallel to OP are $\frac{q}{\text{PM}^2} \cdot \frac{\text{RS}}{\text{RM}} = \frac{q}{\text{PM}^2} \cdot \frac{\text{OP}}{\text{PM}} = \frac{qd}{\text{PM}^3}$, and

$$q \cdot \frac{\text{P'M}}{\text{PM}} \cdot \frac{1}{\text{P'M}^2} \cdot \frac{\text{OP'}}{\text{P'M}} = \frac{q}{\text{PM} \, . \, \text{P'M}} \cdot \frac{\text{OM}}{\text{PM}} = \frac{q}{\text{PM}^2} \cdot \frac{r}{\text{P'M}} = \frac{qd}{\text{PM}^3}$$

respectively, and are therefore equal and opposite. The resultant is therefore along the radius, and the intensity is

$$\mathrm{E}=\frac{qr}{\mathrm{PM}^3}-\frac{qd^2}{r\,.\,\mathrm{PM}^3}=\frac{qr}{\mathrm{PM}^3}\left(1-\frac{d^2}{r^2}\right)$$

But by Coulomb's theorem (p. 126), $\mathrm{E}=4\pi\sigma$,

$$\therefore\ \sigma=\frac{qr}{4\pi\,.\,\mathrm{PM}^3}\left(1-\frac{d^2}{r^2}\right).$$

We therefore see that the surface density of charge on a sphere when a point charge is brought into its neighbourhood ceases to be uniform, but the attraction between the two may nevertheless be calculated by means of the method of electrical images.

CHAPTER V

ELECTROSTATICS (*continued*)

Measurements (Capacity)

We have seen that for any given dielectric, the ratio of the induction to the electric intensity is called the dielectric constant or specific inductive capacity. The absolute constancy of this quantity has only been established for media such as the gases, for which the value of k is very nearly unity: in other cases its value depends upon the time for which the field is applied. Thus, if its value be deduced from measurements made with very rapidly alternating fields, the value of k is found to diminish as the alternations become more rapid. For the present, however, we shall deal with k as though it had a definite and constant value for each substance.

Sphere.—Consider an insulated sphere whose potential is originally zero; when a charge q is placed upon it, the potential at the surface of the sphere becomes $\frac{q}{a}$, where a is the radius of the sphere, provided that the charge is uniformly distributed over it, and that the sphere is surrounded by air. The potential is proportional to the charge, and the ratio of one to the other is called the *capacity*.

Thus for the sphere, $V=\frac{q}{a}$, or $\frac{q}{V}=a$, and the capacity is equal to the radius.

In any case the capacity of a conductor is the ratio of the charge placed upon it to the resulting change in potential, or is the amount of charge which will raise the potential by unity.

Concentric Spheres.—The capacity of a sphere of radius a surrounded by a concentric sphere of radius b maintained at zero potential may now be found (Fig. 147).

The inner surface of b, being conducting, is an equipotential surface, and therefore when a charge $+q$ is placed on a, there will be an equal and opposite charge $-q$ situated upon the inner side of b (see p. 127). Then potential of a due to its own charge is

$+\frac{q}{a}$, and the potential throughout the space inside b due to the charge on the inner surface of b is $-\frac{q}{b}$

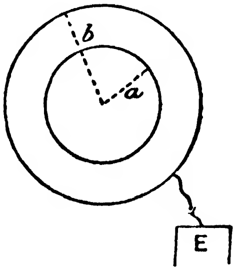


FIG. 147.

$\therefore$ resulting potential of $a=\frac{q}{a}-\frac{q}{b}=\mathrm{V}$.

But capacity$=\frac{q}{\mathrm{V}}$,

$$\therefore \mathrm{C}=\frac{1}{\frac{1}{a}-\frac{1}{b}}=\frac{ab}{b-a}.$$

This is necessarily greater than a, and therefore the capacity is increased by the presence of b.

For practical purposes the zero of potential is taken as that of the earth. Since potential is only recognizable by its differences, we may take that of the earth as zero without affecting our calculations. Thus we say that the sphere b is "earthed."

Cylindrical Condenser.—The capacity per unit length of a cylinder surrounded by an earthed coaxial cylinder is determined by finding the difference of potential between the inner and outer coatings when the charge per unit length of the inner one is $+q$.

The electrical intensity at a distance r from the axis (Fig. 148) due to the charge upon a is $\frac{2q}{r}$ (see p. 125). Therefore p.d. between inner and outer cylinder,

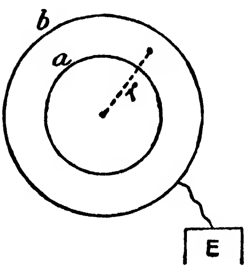


FIG. 148.

$$\mathrm{V}_a-\mathrm{V}_b=-\int_b^a\frac{2q}{r}dr$$

$$=-2q\Big[\log_\epsilon r\Big]_b^a$$

$$=2q\log_\epsilon\frac{b}{a}.$$

But if b is earthed, $\mathrm{V}_b=0$,

$$\therefore \text{capacity per unit length}=\frac{1}{2\log_\epsilon\frac{b}{a}}.$$

Parallel Plate Condenser.—The capacity per unit area of an insulated plate at a distance t from an earthed plate parallel to it may be found by giving the insulated plate a charge of surface density $+\sigma$. This is the surface density of the charge which

faces the earthed plate, and will, unless there are conductors brought near the remote side of A (Fig. 149), be practically the whole of the charge upon A. If such conductors are brought near A, of course the problem is changed, but we are only concerned here with the charge on the side facing B. Electric intensity in the space between A and B is $4\pi\sigma$ (p. 126).

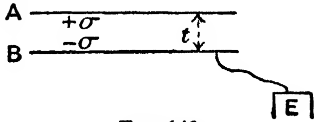


FIG. 149.

$$\therefore\ V_A - V_B = \int_B^A 4\pi\sigma \,.\, dt$$
$$= 4\pi\sigma t,$$

$$\therefore \text{ capacity of unit area of A} = \frac{\sigma}{4\pi\sigma t} = \frac{1}{4\pi t}.$$

Effect of Dielectric on Capacity.—The effect of a dielectric other than air separating the conductors is to increase the capacity in all cases.

Thus for the cylindrical condenser, $E = \dfrac{2q}{kr}$,

$$V_a - V_b = -\int_b^a \frac{2q}{kr} dr,$$
$$V_a = \frac{2q}{k} \log_\epsilon \frac{b}{a},$$
$$\text{and capacity } = \frac{k}{2\log_\epsilon \dfrac{b}{a}}.$$

For the plates, field between A and B $= \dfrac{4\pi\sigma}{k}$ (Fig. 149),

$$\therefore\ V_A - V_B = \frac{4\pi\sigma t}{k},$$
$$\text{and capacity } = \frac{k \times \text{area}}{4\pi t}.$$

In fact, in any case, since the electric intensity is everywhere diminished in the ratio $1 : k$, the potential is diminished in the same ratio, and the charge required to bring the potential back to that for an air condenser must be increased in the ratio $k : 1$. It was owing to the increase in capacity caused by various dielectrics that the effect of the medium was first noticed. Faraday measured the dielectric constant or, as he called it, the *specific inductive capacity* of a number of substances by comparing the capacity of two similar condensers, one having the substance and the other air as dielectric.

Condensers.—The Leyden jar (Fig. 150) is a form of condenser used in electrostatic experiments. It is a glass jar coated inside and outside for about two-thirds of its depth with tinfoil, the remaining glass surface being coated with shellac varnish to improve the insulation. The outer coating is usually sufficiently well earthed by standing on an ordinary table, and contact with the inner coating is made through a metallic knob and stand. If A be the area of the inner coating and t the thickness of the glass, the capacity is about $\frac{6A}{4\pi t}$, the dielectric constant of the glass being about 6.

Many *standard condensers* are constructed of layers of tinfoil separated by sheets of mica, the alternate layers of tinfoil being

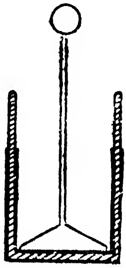


FIG. 150.

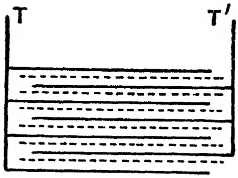


FIG. 151.

connected respectively to the terminals T and T′ (Fig. 151). The mica being an extremely good insulator and at the same time very thin, a great capacity can be obtained without the necessity for great bulk. The dielectric constant of mica is about 6. For convenience in use, condensers are made up in a manner similar to that of resistance boxes, but it must be remembered that to add the capacities the condensers must be placed in parallel, not in series. With the plugs as shown in Fig. 152 the capacity 0·75 microfarad is being used.

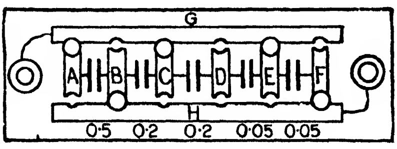


FIG. 152.

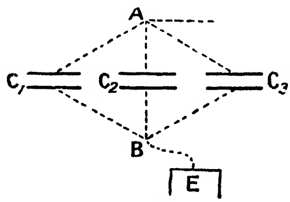


FIG. 153.

Condensers in Parallel.—Thus in Fig. 153, C_1, C_2, C_3 are three condensers connected in parallel between A and B, and the same p.d. exists between the terminals of the three. Let p.d. $=V$.

Then if q_1, q_2 and q_3 be the charges upon each,

$$q_1=C_1V,\ q_2=C_2V,\ q_3=C_3V,$$

and,

$$q_1+q_2+q_3=(C_1+C_2+C_3)V.$$

$$\text{But, total capacity } C=\frac{\text{total charge}}{V}$$

$$=\frac{q_1+q_2+q_3}{V},$$

$$=C_1+C_2+C_3.$$

Condensers in Series.—If the condensers are connected in series as in Fig. 154, the combined capacity may be found from the fact that the charges upon the opposite plates of each condenser are equal. Thus if the charge $+q$ is situated upon one plate of C_1, $-q$ is situated on the other. Hence the charge $+q$ has passed to the first plate of C_2, and so on.

Then,

$$V_A-V_E=\frac{q}{C_1},$$

$$V_E-V_F=\frac{q}{C_2},$$

$$V_F-V_B=\frac{q}{C_3},$$

$$\therefore\ V_A-V_B=q\left(\frac{1}{C_1}+\frac{1}{C_2}+\frac{1}{C_3}\right).$$

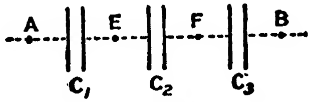


FIG. 154.

But, $\dfrac{q}{V_A-V_B}$ is the combined capacity C,

$$\therefore\ \frac{1}{C}=\frac{1}{C_1}+\frac{1}{C_2}+\frac{1}{C_3}.$$

In the case of most condensers used for practical purposes, the capacity has been found by comparison with a standard, whose value can be calculated from its geometrical form. The simplest case is that of a sphere whose capacity is equal to its radius, but a sphere would have to be of such great size to have a sufficient capacity for practical purposes that this consideration alone would prevent its use. But in addition, we have the fact that the sphere must be at an infinite distance from all other bodies for the surface density of charge to be uniform and the field everywhere radial. A pair of concentric spheres would get over both these difficulties, but then we meet with the objection that it is difficult to construct the spheres and to arrange them to be concentric, and the insulation of the inner sphere would also give trouble.

Guard-Ring Condenser.—The first satisfactory standard condenser to be made was the guard-ring condenser of Lord Kelvin.

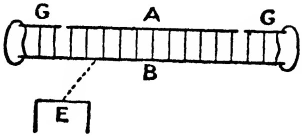


FIG. 155.

In the case of the parallel plate condenser, an uncertainty in the effective area of the insulated plate arises from the fact that the field near the edge of the plate is not uniform (see Fig. 155). Lord Kelvin got over this difficulty by making the insulated plate circular and surrounding it by a guard-ring, so that the irregularity of the field does not occur at the edge of A, but at that of the guard-ring G. There is still a slight irregularity in the field where the gap between A and G occurs, but if this gap is not very wide, Kelvin found that the effective area of the plate is the arithmetical mean of the area of the plate A and the circular hole in G. Calling this effective area A,

$$\text{Capacity} = \frac{kA}{4\pi t}.$$

Fig. 156 illustrates the guard-ring condenser,[1] the left-hand half of the figure being in section. *h*, the guard-ring, and *k*, the

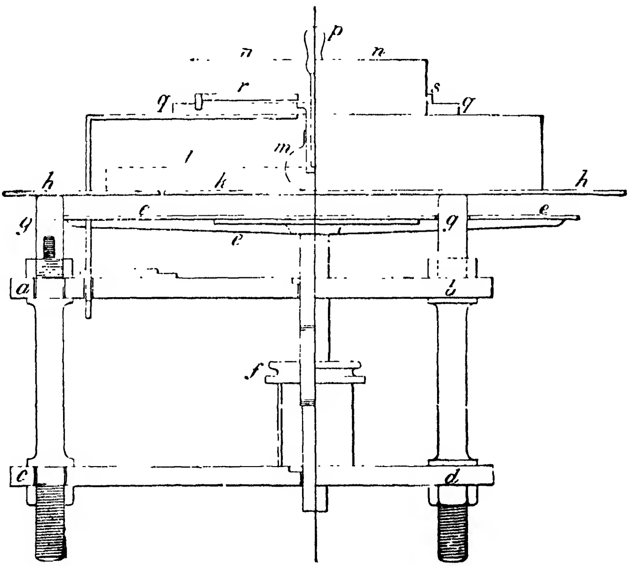


FIG. 156.

plate, can be insulated or connected together, and the parallel earthed plate *e* can be raised or lowered by means of the micrometer screw *f*. The distance apart of the plates is measured by

[1] J. Hopkinson, *Phil. Trans.*, 169, p. 17. 1878.

bringing e into contact with k and h. The reading of the micrometer screw is observed and e is then lowered by turning the screw, and the reading again taken. On subtracting one micrometer reading from the other, the distance t is obtained. In using the instrument, e is earthed, and h and k connected together and charged. On then earthing h, the charge upon k remains and has the value corresponding to the capacity $\frac{kA}{4\pi t}$.

Variable Condensers.—One convenient form is the sliding or cylindrical condenser, also due to Lord Kelvin. A and C (Fig. 157) are two coaxial cylinders of the same diameter, with a small

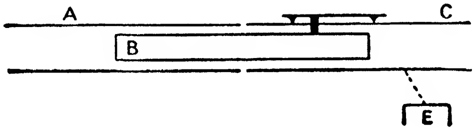


FIG. 157.

gap separating them. The smaller cylinder B is coaxial with the other two and is carried by a support which slides upon C, so that the length lying within A can be varied and measured. A is insulated and B and C are earthed. Then, if B is caused to slide in or out of A by the distance l, the change in capacity of A is $\dfrac{kl}{2\log_e \frac{a}{b}}$. A is surrounded by an earthed metallic tube to prevent its capacity from being varied by the movement of conductors in its neighbourhood. This condenser has not a capacity of known absolute value, but its change in capacity for any movement of B is known from its dimensions. If a scale of lengths be attached to the slider, and the absolute capacity be determined experimentally for one position, that for other positions of the scale will then be known.

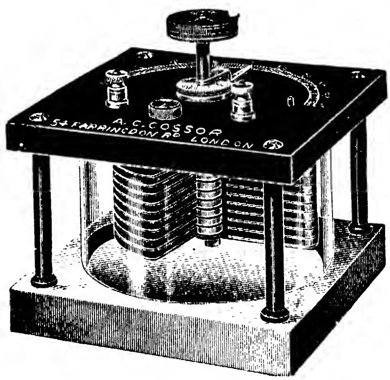


FIG. 158.

Another useful variable condenser made by Mr. A. C. Cossor, is shown in Fig. 158. A number of semicircular plates are insulated, and between them and parallel to them are a number of similar plates mounted upon an axle, so that they may be rotated to occupy any position from lying entirely within

or entirely external to the fixed plates. This is really a multiple parallel plate condenser, and the relative position of the plates may be determined by means of a pointer moving over a circular scale. The condenser is not an absolute one, and the scale must be calibrated by experimental comparison with known standard condensers. It is the type from which those now commonly used in wireless work are derived.

Electrometers.—For the comparison of capacities, some measurer of potential is necessary. The electromagnetic voltmeter described in Chapter III is, of course, useless, since its reading depends upon the existence of a current, which is exactly what must be avoided in dealing with electrostatic charges; in fact, the difficulties met with are largely due to faulty insulation, which allows minute currents to flow, and hence the charges to leak away.

Faraday used the gold-leaf electroscope as an electrometer, but its low sensitiveness and the uncertainty in the value of its readings restrict its use. The gold-leaf electroscope has been so modified in form that its use as an electrometer has been greatly extended (see Chap. XV).

Attracted Disc Electrometer. — More exact electrometers, depending upon a measurement of the attraction between two conductors, maintained at different potentials, have been designed, and after many modifications, the instrument took the form of the *Attracted Disc Electrometer*, which, in the case of Lord Kelvin's pattern, is an "absolute" instrument, the potential difference being found in terms of a force, a length and an area. The arrangement is shown diagrammatically in Fig. 159. A is the attracted plate, which is carried by a spring S, and situated in the plane of the guard-ring C, just as in the case of the guard-ring condenser. A is maintained at a constant potential, and B, which can be raised or lowered by means of the micrometer screw M, is brought in turn into contact with the bodies, the difference of potential between which it is required to determine.

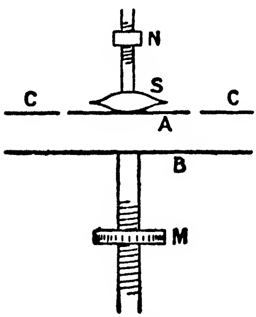


FIG. 159.

To calculate the attraction between A and B, let their potentials be V_a and V_b and their distance apart t. The electric intensity E in the space between them is therefore $\frac{V_a - V_b}{t}$.

$$E = \frac{V_a - V_b}{t}.$$

Also from p. 131, we know that the force per square centimetre of either is $\frac{E^2}{8\pi}$, so that if A be the area of the plate A, the total force upon it is—

$$\frac{E^2}{8\pi}A=F,$$

$$\therefore F=\frac{(V_a-V_b)^2A}{8\pi t^2},$$

$$\text{or, } V_a-V_b=t\sqrt{\frac{8\pi F}{A}}.$$

In using the instrument, the plates are, to begin with, all earthed, and a small known weight m is placed upon A. This depresses it below the plane of C, and it is raised by means of the screw N until it again comes into the plane of C. The small weight is then removed, when, of course, the spring S pulls it above C, but if subsequently the attraction between A and B pulls A again into the plane of C, we know that the force acting on it is mg dynes, where g is the acceleration of gravity. A and C are now insulated and charged to a potential which is kept constant by means of a small electrical machine, the prototype of the Wimshurst machine, called the Kelvin replenisher, the constancy being observed by means of a second guard-ring condenser with fixed spring control and fixed position of earthed plate. B is then connected to the first of the points the p.d. between which we wish to measure, and its position adjusted by means of M until A is in the plane of C. The force on A is then mg, and if V_1 be the potential of the first point, and the micrometer M is read—

$$V_a-V_1=t_1\sqrt{\frac{8\pi mg}{A}}.$$

B is now connected to the second point, whose potential is V_2, and the adjustment again made.

$$V_a-V_2=t_2\sqrt{\frac{8\pi mg}{A}}.$$

Therefore, $$V_2-V_1=(t_1-t_2)\sqrt{\frac{8\pi mg}{A}}.$$

Thus V_2-V_1 is known in terms of the difference of the two micrometer readings, the area of the plate A, and the weight of a small mass m. It is not necessary to know either the actual

distance between A and B or the potential of A, but the latter must remain constant.[1]

Quadrant Electrometer.—The Quadrant Electrometer bears a certain resemblance to the galvanometer, in that both give deflections proportional to the quantity to be measured, the deflection in each case depending upon a number of constants of such an uncertain value that the instrument is used for purposes of comparison only. The quadrant electrometer, like that of the attracted disc type, is evolved from others of simpler type and achieved its first satisfactory form at the hands of Lord Kelvin.

The Fig. 160 shows a typical arrangement of the instrument. Four hollow quadrants are connected in pairs AA and BB. A paddle-shaped conductor C, frequently called the "needle" from analogy with the galvanometer needle, is maintained at a high (or low) potential and is therefore positively (or negatively) charged. When A and B are at the same potential, the needle

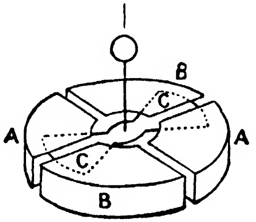


FIG. 160.

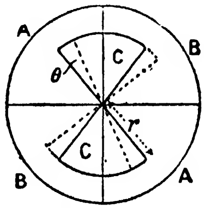


FIG. 161.

hangs symmetrically between them, but on establishing a difference of potential between A and B, the field produces a couple which rotates C until the control brings it to rest. The deflection is approximately proportional to the potential difference between A and B. To prove this, imagine that the needle is charged to a high potential V_C, and the quadrants to potentials V_A and V_B, where $V_A > V_B$. Let these potentials be maintained by connection with sources of supply such as cells. Consider the needle to be held for a moment in its zero position and then released. It will then be deflected towards the B quadrants; that is, down the grade of potential. The requisite energy is all drawn from the source of supply, and is used in twisting the suspension and in increasing the potential electrical energy associated with the quadrants and needle. Equilibrium is attained when the deflecting couple is equal to the restoring couple due to the suspension fibre.

Consider a deflection θ (Fig. 161); then an area of needle equal

[1] For a detailed account of the construction of the absolute electrometer and its use, the student may consult "Absolute Measurements in Electricity and Magnetism," by A. Gray.

to $\pi r^2\theta/\pi = r^2\theta$ has been transferred from the A quadrants to the B quadrants, and, remembering that there are two faces to the needle, the effective area transferred is $2r^2\theta$. This is equivalent to diminishing the capacity of the A–C condenser by an amount $2r^2\theta/4\pi t = r^2\theta/2\pi t$, and increasing that of the B–C condenser by the same amount, t being the thickness of air space between needle and quadrants. Thus an amount of charge $r^2\theta(V_C - V_A)/2\pi t$ has been lost from the A–C condenser, and the corresponding loss of electrical energy is $(V_C - V_A)r^2\theta(V_C - V_A)/2\pi t = r^2\theta(V_C - V_A)^2/2\pi t$, since the potentials have remained constant throughout. Similarly the amount of energy gained from the sources of supply on the B–C side is $r^2\theta(V_C - V_B)^2/2\pi t$. Thus the total energy supplied to the electrometer from the sources is—

$$\frac{r^2\theta}{2\pi t}\{(V_C - V_B)^2 - (V_C - V_A)^2\} = \frac{r^2\theta}{\pi t}(V_A - V_B)\left(V_C - \frac{V_A + V_B}{2}\right).$$

Again, the potential energy of the charges residing on the A–C condenser has been reduced by $\frac{1}{2}$(capacity)(p.d.)$^2 = r^2\theta(V_C - V_A)^2/4\pi t$, and the gain on the B–C side is $r^2\theta(V_C - V_B)^2/4\pi t$. Therefore the gain of potential energy of the charges residing on the electrometer is $\frac{r^2\theta}{2\pi t}(V_A - V_B)\left(V_C - \frac{V_A + V_B}{2}\right)$, and the balance of work done by the source of supply over the potential energy gained by the charges is $\frac{r^2\theta}{2\pi t}(V_A - V_B)\left(V_C - \frac{V_A + V_B}{2}\right)$, which is proportional to θ. Hence the deflecting couple is constant and equal to $\frac{r^2}{2\pi t}(V_A - V_B)\left(V_C - \frac{V_A + V_B}{2}\right)$.

If the restoring couple of the suspension fibre for one radian twist is c, that for deflection θ is $c\theta$, and for equilibrium the two couples are equal,

or, $$\theta = \frac{r^2}{2\pi ct}(V_A - V_B)\left(V_C - \frac{V_A + V_B}{2}\right).$$

We see, therefore, that the deflection is proportional to $\frac{r^2}{2\pi ct}$, which is a constant, to the difference of potential between A and B, and to the difference between the potential of C and the average potential of A and B. The last term is very nearly constant if V_C is great and V_A and V_B change very little during the experiment, and we may then say that—

$$\theta = K(V_A - V_B)$$

where K is a constant.

The shape of the needle is immaterial, provided that the change in area within each pair of quadrants is proportional to the

deflection, which condition is fulfilled when the outer edge of the needle is circular, and the radial edges of the needle lie well within the quadrants.

In the method of use described above, the conductors A, B and C are all at different potentials, and the instrument is said to be used *heterostatically*. It may also be used *idiostatically* by connecting C to one pair of quadrants, say A. Then $V_A=V_C$, and the equation for the deflection becomes—

$$\theta=\frac{r^2}{4\pi ct}(V_A-V_B)^2$$
$$=K'(V_A-V_B)^2.$$

In this case the constant K′ is much smaller than K, since V_C is very great ($2K'V_C=K$). Hence for a given deflection, V_A-V_B will generally be greater when the use is idiostatic than when heterostatic. The range of usefulness of the instrument is therefore considerably extended. It should be noted, however, that when the method is idiostatic, the deflection is proportional to the square of the p.d. to be measured. This is for some purposes inconvenient, but on the other hand it has the great advantage that the deflection is in the same direction whether the p.d. is positive or negative, and the electrometer will therefore give a deflection with an alternating potential difference. This point will be dealt with in the chapter on alternating currents.

In the Kelvin instrument V_C is maintained constant by means of a replenisher and a trap-door indicator, as in the case of the guard-ring electrometer, and further, the constancy is assisted by placing C in connection with the insulated coating of a condenser of large capacity. Also the control is produced by means of a bifilar suspension.[1]

The type of quadrant electrometer most commonly used at the present time was designed by F. Dolezalek, and is shown in Fig. 162. The quadrants are of brass, and are mounted on amber pillars A to ensure good insulation. The needle N is of paper, thinly coated with metal, or in the latest instruments it is of thin aluminium. It is suspended by a quartz fibre which is made conducting by dipping it into a strong solution of calcium chloride, which, being very hygroscopic, maintains the surface of the fibre sufficiently conducting to keep the needle charged by means of a battery. The quartz fibre produces such a feeble control that a high sensitiveness is obtained without the employment of very high potential for C. With a suitable fibre and a p.d. of 50 to 100 volts between the needle and earth, one volt will cause a deflection of 200 to 400 millimetres with a scale distance of a

[1] For further description of the Kelvin quadrant electrometer the student is referred to the pamphlet by Messrs. Kelvin and White of Glasgow.

metre from the mirror. The deflection is usually observed by means of a lamp and scale, as in the case of the reflecting galvanometer. If the absolute values for the deflections are required, the scale must be calibrated by some known source of p.d., as for example a standard cell. If an insulating quartz fibre be used for the suspension, this will enable the needle to keep the charge for some time. The needle may always be recharged by means of the contact maker K.

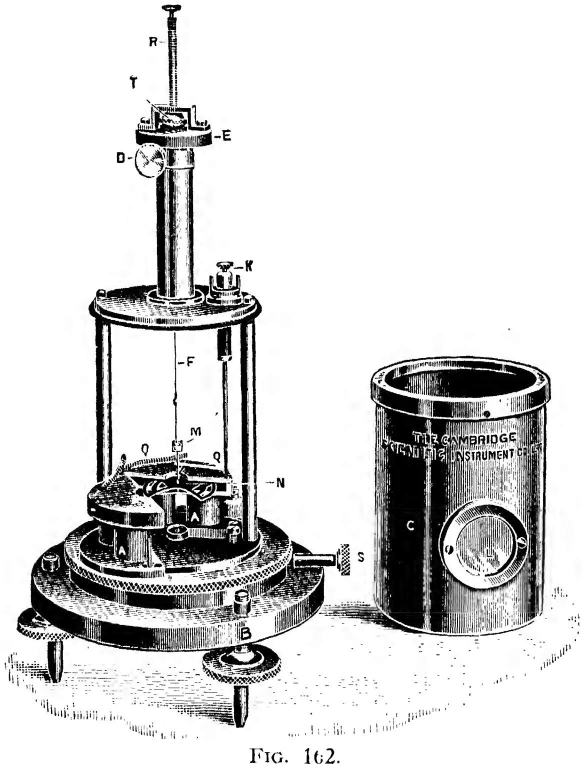


FIG. 162.

To control the sensitiveness of the quadrant electrometer, A. H. and K. T. Compton[1] used a slightly tilted needle and had the level of one pair of quadrants slightly above that of the other. Small changes in this level alter the sensitivity and up to 60,000 mm. per volt is attainable.

The portable Lindemann electrometer[2] has a modified arrangement, with a taut quartz suspension controlling the needle, which is a stiff metal-coated glass fibre. The tip of the needle is observed

[1] A. H. Compton and K. T. Compton. *Phys. Rev.*, XIV (2nd), p. 85. 1919.
[2] F. A. and A. F. Lindemann and T. C. Keeley, *Phil. Mag.*, **47**, p. 577. 1924.

with a microscope or an image of it is projected optically. The capacity and period are both very small.

Several forms of *Electrostatic Voltmeter* operate on similar principles to the quadrant electrometer, but with the needle pivoted. To improve the sensitivity despite the more robust construction, the Kelvin multicellular form may be used, with a large number of alternating quadrants, connected idiostatically. The moving system may be suspended by a metal strip which also supplies the control.

String Electrometer.—A bifilar instrument of moderate sensitiveness and very small capacity is illustrated in Fig. 163. A pair of fine fibres of platinum are fixed to a terminal at A and a quartz bow B, which exerts a tension on them. The conductors E and F are earthed. When the fibres are at a potential, not zero, they stand apart, the distance between them being measured by means of a microscope. The instrument must be calibrated, either by means of standard cells or an electrometer whose readings for given p.d.'s are known. A modified form consists of a single fibre between two plates, one of which is earthed and the other charged. The instrument has extremely small capacity and its readings are almost instantaneous. Its principle is really that of the gold-leaf electroscope.

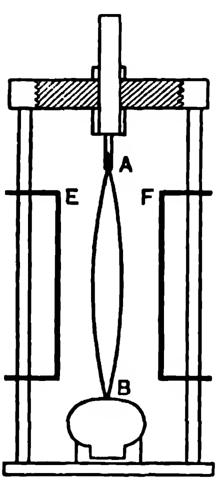


FIG. 163.

Comparison of Capacities.—On p. 149 we saw that the capacity of any conductor is directly proportional to the dielectric constant of the medium in which it is situated, and therefore the measurement of the constant of any dielectric generally devolves upon the comparison of the capacities of two condensers, one with and the other without the given medium as dielectric.

Faraday's Method.—The first determination of dielectric constant was made by Faraday. He constructed two spherical condensers of the pattern shown in Fig. 164, as nearly as possible alike, and tested their equality in capacity by charging the inner sphere of one of them and then sharing its charge with the other, the outer spheres being earthed. He found that the potential fell to half, on the sharing of the charge taking place. The lower half of one of them, B, was then filled with shellac, and the other one, A, which had only air between the spheres, was given a charge. On sharing A's charge between the two, the potential did not fall to exactly half, showing that the capacities were no longer equal. Thus, if V_1 is the original potential of A, and V_2 the final common potential, the original potential of B being zero, calling C_a and C_b the capacities of A and B—

The charge passing from A to B, that is $q=C_a(V_1-V_2)=C_bV_2$,

$$\therefore \frac{C_b}{C_a}=\frac{V_1-V_2}{V_2}$$

In this way it was found that $C_b=1\cdot50C_a$. Remembering that only half of B is filled with shellac of dielectric constant k, and that without shellac the capacities are equal,

$$C_b=k\,.\,\frac{C_a}{2}+\frac{C_a}{2}=(k+1)\frac{C_a}{2},$$

$$\therefore\ 1\cdot50=\frac{(k+1)}{2},$$

which gave $k=2$ for shellac.

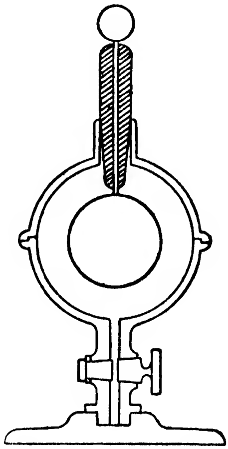


FIG. 164.

In a similar manner Faraday found the dielectric constant of sulphur to be 2·2.

A considerable improvement in the comparison of capacities by the above method may be made by using the electrometer instead of the electroscope, but in this case the capacity of the electrometer itself may be appreciable, and should be taken into account.

If A be a standard condenser, it is first charged by depressing the key p (Fig. 165), and the electrometer deflection noted; let it be θ_1. The key p is then opened and q closed, and the new deflection θ_2 noted. Since the potentials in the two cases are proportional to θ_1 and θ_2, and the capacity of the electrometer (C) must be added to that of A, we have as before the relation,

$$\frac{C_b}{C_a+C}=\frac{\theta_1-\theta_2}{\theta_2}.$$

C may be found by charging A and the electrometer and noting the deflection θ_3. The electrometer is then insulated, discharged, and again connected to A and the deflection θ_4 noted.

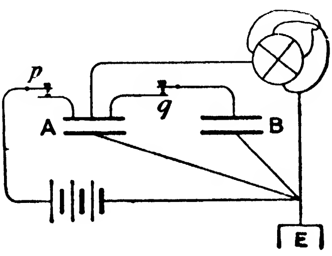


FIG 165.

The charge which passed from A to the electrometer being q,

$$q=C_a(V_3-V_4)=CV_4$$

$$\frac{C}{C_a}=\frac{V_3-V_4}{V_4}=\frac{\theta_3-\theta_4}{\theta_4}.$$

If C is very small, the experiment does not give an accurate determination, owing to the smallness of $\theta_3-\theta_4$; but it should be noted that the smaller

that C is, the less is the importance of knowing its value accurately.

If one of the capacities, say C_a, be very small in comparison with the other, it may be charged a number of times from the other, provided that it is discharged between the successive chargings. When C_a is first connected to C_b the charge remaining on C_b is $Q\frac{C_b}{C_a+C_b}$, where Q is the initial charge upon C_b. If C_a be then discharged and again connected to C_b, the charge upon the latter falls to

$$Q\left(\frac{C_b}{C_a+C_b}\right)\left(\frac{C_b}{C_a+C_b}\right)=Q\left(\frac{C_b}{C_a+C_b}\right)^2.$$

For n charges and discharges, the charge upon C_b falls to

$$Q\left(\frac{C_b}{C_a+C_b}\right)^n,$$

and the potential falls proportionately. Thus, if V_1 is the initial potential of C_b, and V_n the potential after n sharings—

$$\frac{V_n}{V_1}=\left(\frac{C_b}{C_a+C_b}\right)^n.$$

V_n and V_1 being proportional to the electrometer deflections and C_b and n being known, C_a may be calculated. Fig. 166 shows how the repeated sharing may be performed by means of the key K.

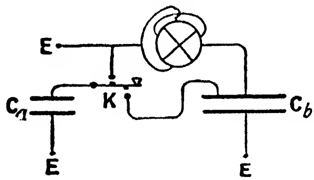


FIG. 166.

Kelvin's Methods.—The following null method of comparing capacities is due to Lord Kelvin. Four condensers are required, of which one must be variable, and these are arranged as in Fig. 167, the points A and B being joined to some source of electromotive force. Then if the points E and F remain at the same potential, the electrometer needle will be undisturbed, but if not there will be a deflection. If a balance is not obtained, the capacity of the variable condenser is altered and the experiment repeated. When a balance is reached the p.d. between A and E, *i.e.* $\frac{q_1}{C_1}$, is equal to that between A and F, *i.e.* $\frac{q_2}{C_3}$.

$$\therefore \frac{q_1}{C_1}=\frac{q_2}{C_3}.$$

Similarly, $$\frac{q_1}{C_2}=\frac{q_2}{C_4},$$

$\therefore$ from the two equations, $$\frac{C_1}{C_2}=\frac{C_3}{C_4}.$$

A second null method, also due to Lord Kelvin, is frequently employed for comparing capacities, and it has the great advantage over the previous method, that only one standard condenser is required, and this a constant one, the balancing being effected by varying the resistances in two boxes, R_1 and R_2 (Fig. 168). When the keys 1 and 2 are depressed, the condenser C_1 is charged by the p.d. between the ends of the resistance R_1, the value of which is iR_1, where i is the current produced by the battery in the circuit R_1R_2. A and B are earthed. Hence the positive and negative plates of C_1 have charges $+iR_1C_1$ and $-iR_1C_1$ respectively.

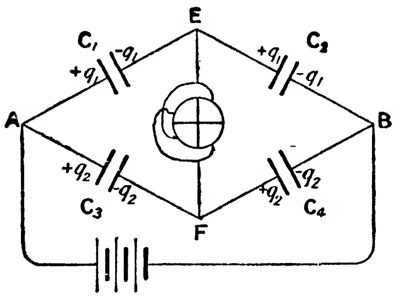


Fig. 167.

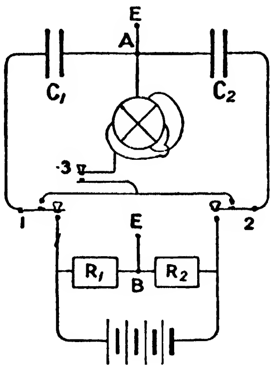


Fig. 168.

Similarly, the charges on C_2 are $+iR_2C_2$ and $-iR_2C_2$. On releasing the keys, the positive plate of C_1 is connected to the negative plate of C_2, and the charge $+i(R_1C_1-R_2C_2)$ will remain. If, now, the key 3 be closed, there will be a deflection unless the remaining charge is zero, in which case $R_1C_1=R_2C_2$. The test is made with various resistances R_1 and R_2, until this condition is fulfilled. This is known as the method of mixtures, and the electrometer may be replaced by a sensitive galvanometer, in which case a transient current flows on closing the key 3, when the balance is not perfect.

Gott's Method.—This is a modification of the last. On closing the key 1, D and F (Fig. 169) will in general have different potentials, and on closing key 2, a deflection will be produced. If, however, the resistances be adjusted so that there is no deflection,

the p.d. between B and D is equal to that between B and F. If, then, E is the p.d. between B and earth

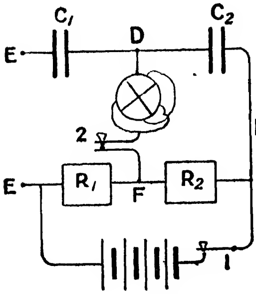


FIG. 169.

$$\text{p.d. between B and F}=E\cdot\frac{R_2}{R_1+R_2},$$

$$\text{p.d. between B and D}=\frac{1}{C_2}\cdot\frac{E}{\frac{1}{C_1}+\frac{1}{C_2}},$$

since $\frac{1}{C_1}+\frac{1}{C_2}$ is the reciprocal of the combined capacity of the condensers in series, and $\frac{E}{\frac{1}{C_1}+\frac{1}{C_2}}$ is therefore the charge on each plate of the condensers (see p. 151).

$$\therefore E\frac{R_2}{R_1+R_2}=\frac{1}{C_2}\cdot\frac{E}{\frac{1}{C_1}+\frac{1}{C_2}},$$

$$\frac{R_2}{R_1+R_2}=\frac{C_1}{C_1+C_2}$$

$$\therefore R_1C_1=R_2C_2.$$

Further methods for measuring capacities will be described in later chapters (VIII, X, XI, XII).

Determination of Dielectric Constant.—The dielectric constant of several solid substances was measured by Boltzmann [1] in 1873 by a method somewhat resembling that of Faraday. A parallel plate condenser is used, between the plates of which could be inserted a slab of the substance. The capacity is compared with that of the electrometer and a fixed condenser together, by means of the method of sharing the charges, both with air between the plates, and when the slab together with an air space separates the plates. In the latter case one plate of the condenser could be moved away from the other until the capacity is restored to its original amount, and the dielectric constant is then known in terms of the thickness of the slab and the displacement of the plate. Thus, if t be the distance apart of the plates and d the thickness of the slab (Fig. 170), ϕ being the electric induction, which is uniform (see p. 152), since the field is everywhere normal to the slab; the electric intensity in

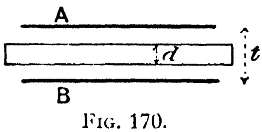


FIG. 170.

[1] L. Boltzmann, *Wien. Akad. Sitzungsber.*, **67** (2), p. 17. 1873.

the air space being E, that in the slab is $\frac{E}{k}$, and if V_a and V_b are the potentials of A and B,

$$V_a - V_b = E(t-d) + \frac{E}{k}d$$
$$= E\left(t - d + \frac{d}{k}\right).$$

But, $\quad E = 4\pi\sigma,$

where σ is surface density of charge on plates,

$\therefore$ capacity per unit area of plate $= \dfrac{\sigma}{4\pi\sigma\left(t-d+\frac{d}{k}\right)}$,

and, capacity of area A $= \dfrac{A}{4\pi\left\{t-d\left(1-\frac{1}{k}\right)\right\}}$

Thus the effect of introducing the slab of thickness d is the same as would be produced by diminishing t by the amount $d\left(1-\frac{1}{k}\right)$, without introducing the slab, so that if t be increased by this amount when the slab is in, the capacity will again be brought to the value it had without the slab. Calling this displacement h,

$$h = d\left(1 - \frac{1}{k}\right),$$

$$\text{or, } k = \frac{d}{d-h}.$$

The values of k found for sulphur, ebonite and paraffin were respectively 3·84, 3·15 and 2·32.

The attracted disc electrometer may be used to measure the quantity h of the last equation. As in Fig. 159, we see that the force per square centimetre of A is $\frac{E^2}{8\pi}$, and the total force $F = \frac{AE^2}{8\pi}$

$$\text{But, } E = \frac{V_a - V_b}{t - d\left(1-\frac{1}{k}\right)},$$

$$\therefore F = \frac{A}{8\pi}\left\{\frac{V_a - V_b}{t-d\left(1-\frac{1}{k}\right)}\right\}^2,$$

$$\text{or, } V_a - V_b = \left\{t - d\left(1-\frac{1}{k}\right)\right\}\sqrt{\frac{8\pi F}{A}}.$$

Thus, if the p.d. between the plates is maintained constant and the slab introduced, F increases and the charged plate is pulled down. On then lowering the earthed plate by the amount $h=d\left(1-\frac{1}{k}\right)$ the charged plate will return to its original position.

Electric Absorption.—The earlier measurements of the dielectric constant were all subject to error, owing to the fact that media other than gases do not instantaneously acquire their maximum induction in an electric field. This phenomenon of "*Electric Absorption*" is very similar to that exhibited by various complex substances, such as glass, when subjected to torsional strain. It was noticed by Faraday in conducting his experiment with the spherical condensers (p. 161), that a smaller result is obtained for k when the condenser containing the shellac is charged first and its charge shared with the other, than when the condenser without the shellac is charged first, and the change in the value of k obtained is greater the longer the interval that elapses between the charging of the shellac condenser and the sharing of the charges.

If a Leyden jar be given a charge and its potential be measured by means of an electrometer, it will be found that the potential will fall for some time, but will eventually become constant. On discharging the jar the whole induction in the medium does not disappear immediately; successive discharges, gradually getting smaller, may be obtained. The charge, which does not disappear at the first discharge, has been called the *residual charge*.

This phenomenon renders it important that in making measurements of the dielectric constant, the time for which the charging takes place, and the interval between charge and discharge should be known. In most cases the result obtained on charging and discharging a condenser within half a second is sufficiently constant to be used in defining the capacity of a condenser for practical purposes.

That it is not necessary to consider that there is an actual absorption of electricity to explain this phenomenon was demonstrated by Maxwell,[1] who showed that a composite layer of dielectrics having different dielectric constants and slight conductivities would exhibit effects like those usually called absorption. It is necessary that the ratio of dielectric constant to resistivity should vary from layer to layer; otherwise there would be no phenomenon corresponding to absorption.

When measurements are made with alternating currents the values of k obtained are independent of any absorption phenomenon.

Hopkinson's Method.—Dr. J. Hopkinson [2] used a modification

[1] James Clerk Maxwell, "A Treatise on Electricity and Magnetism." 1873.
[2] J. Hopkinson, *Phil. Trans.*, 169, p. 17. 1878.

of the method of mixtures for finding the capacity of a guard-ring condenser, and employing the slab method (p. 164) found the dielectric constant of various substances. The charging battery (Fig. 171) is earthed at its middle point and the terminals connected to the plate A of the guard-ring condenser and the inner cylinder D of the sliding condenser. These are consequently at equal and opposite potentials, and if the two capacities are equal, the charges received are equal and opposite. On disconnecting the battery and joining A and D, the resulting potential as indicated by the quadrant electrometer will be zero. The sliding condenser is adjusted until this condition is fulfilled. The slab is now introduced between A and B, and the equality

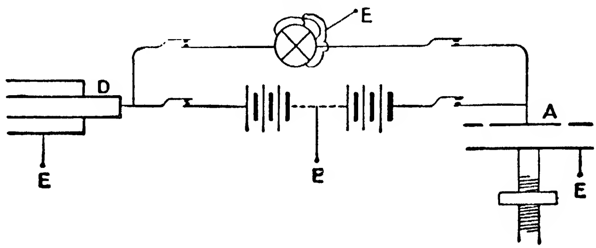


FIG. 171.

re-established by means of the sliding condenser, the change in capacity being therefore known. A special key is used to enable the various connections to be made with great rapidity.

As a result of an extended series of experiments, Hopkinson found that the dielectric constant for glass is constant for times of discharge varying from $\frac{1}{20000}$ to $\frac{1}{4}$ second, for measuring which short periods he used a pendulum for connecting A and D at a known short interval after the charging.

By a modification of the method of sharing charges (p. 161) Hopkinson,[1] using the pendulum make and break, also determined the dielectric constant of certain liquids (Table, p. 169).

Silow's Method.—A method differing entirely from the last has been employed by Silow [2] for determining the dielectric constant of certain liquids. The liquid is made to replace the air in a cylindrical electrometer. The conductors A and B (Fig. 172) are four strips of tinfoil attached to the sides of a cylindrical glass vessel. The needle C is also cylindrical and is made of platinum. It is suspended by a fibre, its deflection being observed in the ordinary way. One pair of conductors, say B, is earthed, the other pair, A, being maintained at steady potential, C also being earthed. The deflection is then proportional to the dielectric constant of the liquid filling the vessel. For, the capacity of a

[1] *Ibid.*, 172, p. 355. 1881.
[2] P. Silow, *Pogg. Ann.*, 156, p. 389. 1875.

condenser being proportional to k, we may introduce this quantity into our calculation for the deflection of the needle of the quadrant electrometer (p. 156), and we then find that, when used idiostatically

$$\theta=\frac{kr^2}{2\pi ct}(V_a-V_b)^2,$$

that is, the deflection is proportional to k.

Cohn and Arons.—Using a modification of Silow's method, Cohn and Arons[1] employed an alternating p.d. for the determination of the dielectric constants of a number of liquids.

An ordinary quadrant electrometer and one constructed on Silow's principle to take the liquid, are connected, as shown in

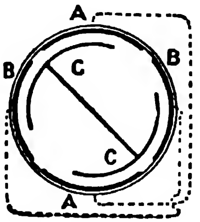


FIG. 172.

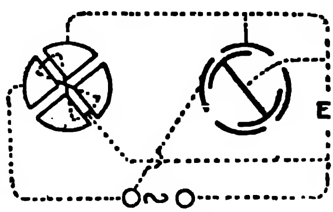


FIG. 173.

Fig. 173. The two electrometers are used idiostatically and are connected in parallel, the p.d. being supplied from an alternating source. Since they are in parallel the p.d. is at every instant the same for the two, and therefore from the equation on p. 158, $\theta=K'(V_a-V_b)^2$, the deflections when both electrometers have air as dielectric are given by,

$$\theta_1=K_1'(V_a-V_b)^2,$$
$$\theta_2=K_2'(V_a-V_b)^2.$$

K_1' and K_2' being the values of $\frac{r^2}{2\pi ct}$ for the electrometers. Since V_a-V_b is the same at every instant for the two,

$$\frac{\theta_1}{\theta_2}=\frac{K_1'}{K_2'},$$

however V_a and V_b may vary. If, now, the liquid be introduced into the Silow electrometer, and two new deflections, θ_1' and θ_2', are simultaneously obtained—

$$\frac{\theta_1'}{\theta_2'}=\frac{K_1'}{kK_2'},$$

$$\therefore k=\frac{\theta_2'}{\theta_1'}\cdot\frac{\theta_1}{\theta_2}.$$

[1] E. Cohn and L. Arons, *Wied. Ann.*, 33, p. 13. 1888.

Other Methods.—In general the dielectric constant of a medium depends on the frequency of alternation of the applied potential, so that it is desirable to make measurements at well-defined frequencies.

In several modern determinations, the test condenser into which the dielectric is to be introduced forms part of an oscillatory circuit (Chapters X and XI). Small differences in frequency between this circuit and an oscillator of fixed frequency are detected by the "beats" produced between them—fluctuations of amplitude of the combined effect due to the two signals coming into and out of phase. The number of beats per second is equal to the difference in the two frequencies and the method is very sensitive to small changes in the frequency of the test circuit and hence to changes in the capacity of the test condenser. When the dielectric is introduced, the original frequency can be restored by altering a calibrated variable condenser in parallel with the test condenser.[1]

For high-frequency measurements with fluids, an alternative method is to utilise the nodes and antinodes which can be detected along a pair of Lecher's wires (p. 459). The wave-length so obtained is inversely proportional to the square root of the dielectric constant, and changes are hence observed when the whole system is immersed in the fluid.[2] Alternatively, oscillations are set up by reflecting electromagnetic waves to and fro in a cylindrical metal cavity and observing changes in the resonant length of the cavity when the cavity is filled with the medium.[3]

Temperature Variation.—The dielectric constant of a given gas depends mainly on its density. Writing $k = 1 + (mp/76)$, where p is the pressure in cm. of mercury, m is found to be independent of temperature for the simpler gases (He and other rare gases, H_2, O_2, N_2) but with some others (*e.g.* NH_3, HCl, H_2O), m falls with rise of temperature.[4] With most liquids k falls with rise of temperature, the effect being especially marked with glycerine, but solids seem mostly to have positive temperature-coefficients.

DIELECTRIC CONSTANTS (k). (From Kaye and Laby's Tables.)

Substance.	k.	Substance.	k.
Crown glass	5–7	Petroleum	2–2·2
Flint glass	7–10	Ethyl alcohol	26·8
Plate glass	6–7	Ether	4·34
Ebonite	2·7–2·9	Water $\lambda = \infty$	81
Sulphur	3·6–4·3	„ $\lambda = 3600$ cm. . . .	3·32
Mica	5·7–7	„ $\lambda = 1200$ cm. . . .	2·79
Paraffin wax . . .	2–2·3	Air (76 cm. pressure, 0° C.) .	1·000594
Shellac	3–3·7	Hydrogen (ditto)	1·000265
Rock salt	5·6	Carbon dioxide (76 cm., 25° C.)	1·000904

[1] G. E. Bell and F. Y. Poynton, *Phil. Mag.* (6), **49**, p. 1065. 1925.
[2] See p. 460. Also R. Bock, *Zeits. f. Phys.*, **31**, p. 534. 1925.
[3] See, for example, L. Essen and K. D. Froome, *Proc. Phys. Soc.* (B), **64**, p. 862. 1951.
[4] M. Jona, *Zeits. f. Phys.*, **20**, p. 14. 1919.

CHAPTER VI

ELECTROLYSIS

Ionic Charge.—In Chapter II we considered Faraday's laws of electrolysis, and saw that the amount of an ion liberated from a solution by an electric current is proportional to the strength of the current and to the time for which it flows ; from this we may now conclude that the amount of the ion liberated is proportional to the amount of electricity which has passed through the electrolyte, since the current itself is the amount of electric charge passing per second. Taking the ampere as the unit of current, the corresponding unit of charge is called the *coulomb,* and is the amount of charge passing when one ampere flows for one second. We can then define the electrochemical equivalent of a substance as the amount liberated by the passage of one coulomb —in fact, Faraday's first law of electrolysis is usually stated to be, " that the amount of deposition is proportional to the quantity of electricity which has passed through the electrolyte." The term ion is used here in the sense originally intended by Faraday, for the substance liberated by the passage of a current. At the present time the word ion has changed in meaning and is applied to the particle, which may be an atom, or group of atoms, which has a positive or negative charge. In this sense ions may exist in a liquid or gas, and on the application of an electric field will travel towards either cathode or anode according to the sign of the charge.

The second law of Faraday states that a given quantity of electricity passing through the electrolyte liberates an amount of substance proportional to its chemical equivalent, and it follows that the amount of a monovalent ion liberated is proportional to its atomic weight, of a divalent ion to half the atomic weight, and so on. Thus 107·88 grammes of silver, 35·46 grammes of chlorine, 62 grammes of NO_3, 31·77 grammes of copper, etc., are each liberated by the same amount of electricity passing through the cell. This amount may conveniently be taken as a unit of quantity of the substance, and is called a *Gramme-equivalent.* Thus a gramme-equivalent of any monovalent substance is a quantity which, measured in grammes, is numerically equal to the atomic weight, and of a divalent substance to half the atomic

weight, etc. The importance of this lies in the fact that a gramme-equivalent of any substance is liberated by the passage of a fixed amount of electricity through the electrolyte. This amount of electricity may easily be found from the electro-chemical equivalent, for it is the amount deposited by the passage of one coulomb ; therefore, taking the electro-chemical equivalent of silver to be 0·0011183 and the atomic weight 107·88—

Charge required to pass in order to liberate one gramme-equivalent of silver $=\frac{107 \cdot 88}{0 \cdot 0011183}=96{,}467$ or $96{,}470$ coulombs.

Since the gramme-equivalent of all monovalent substances contains the same number of atoms, we see that the liberation of each atom from the electrolyte requires the same amount of charge, and again, since the gramme-equivalent of a divalent element has half this number of atoms, all divalent substances require twice this amount of charge for the liberation of an atom.

These facts strongly suggest that the atoms are the carriers of the charges, and that a monovalent atom carries a constant amount of electricity, whatever be its chemical nature, a divalent atom twice that amount, a trivalent atom three times that amount and so on ; and further, that the metallic atoms being liberated at the cathode have a positive charge, and the non-metallic atoms or radicles, since they are liberated at the anode, have a negative charge.

Conduction.—So far we have not made any assumption as to the mechanism of the transport of the atoms with their respective charges from the solution to the electrode, but it becomes of the greatest importance to decide whether the positively and negatively charged atoms forming a molecule of the substance in solution are pulled apart on the application of the electrical field which produces the current, or whether they are wandering about independently of each other in the solution, and are merely subjected to forces, just as any other charged bodies would be, which drive the positively charged atoms down the grade of potential and those negatively charged up the grade of potential.

In future we shall speak of an atom with its associated charge as an ion ; thus in an electrolyte the positive ions are liberated in the neighbourhood of the cathode, and these, on giving up their charges to the electrode, acquire the properties of neutral chemical atoms.

Experience shows that a definite amount of energy must be expended in order to effect the separation of the two ions forming a binary molecule, which energy reappears, usually in the form of heat, upon their recombination. Hence, if the ions in an electrolyte are all in a state of combination to form neutral mole-

cules, we should expect that a certain minimum potential difference would be necessary before any decomposition would occur. But, on the contrary, it is found that any potential difference, however small, will cause *some* current to flow, although the current will soon cease unless the potential difference between the electrodes exceeds a certain amount; about 1·7 volt in the case of acidulated water. From this it might at first sight be concluded that there are a few unattached ions in the electrolyte, and that when all these have been driven to the electrodes the current ceases.

Polarisation.—The real cause of the stoppage of the current, however, is to be sought in the layer of ions which collects upon the electrodes, which layer produces an electromotive force in opposition to that driving the current. The phenomenon is called *polarisation*, and it may be exhibited by immersing two platinum plates in a dilute solution of sulphuric acid, and passing the current by depressing the key K (Fig. 174). On releasing the key, the battery is disconnected, and the galvanometer connected to the plates, when it will be found that a current will flow for a short time. The deposition of hydrogen ions upon the cathode *k* and oxygen upon the anode *a* produces a back electromotive force, and a reverse current flows when the battery is removed and the circuit completed. This back electromotive force causes the reverse current to flow until the collected hydrogen and oxygen ions have been removed.

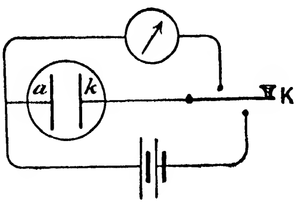


FIG. 174.

We should expect on general grounds that some minimum electromotive force would be required to decompose any substance continuously; for in order to decompose one gramme-molecule of a substance such as water (in this case 18 gm.) energy is required to separate the hydrogen and oxygen ions, the amount of which (68,400 calories per gramme-molecule) may be determined by finding the energy liberated in the form of heat on allowing them to combine. When one gramme-molecule of water is decomposed, two gramme-equivalents of hydrogen are liberated, and therefore $2 \times 96{,}470$ coulombs have passed through the electrolyte. If this passage is caused by an electromotive force equal to E volts, $2 \times 96{,}470 \times E$ joules is the amount of work done, and this must be at least as great as the amount of energy liberated when a gramme-molecule of water is formed. Thus, if no other work is performed by the electromotive force in the cell,

$$2 \times 96{,}470 \times 0{\cdot}239 \times E = 68{,}400,$$

from which $E=1{\cdot}48$ volts, and we cannot think that a less electromotive force can continuously decompose water. For, if this were the case, we could derive more energy from the liberated hydrogen and oxygen by the process of combustion than was used in separating them, and, by a suitable mechanism, we should then have an inexhaustible supply of energy, which contradicts our experience. The actual back electromotive force, opposing the current when platinum plates are used as described above, is about 1·7 volts, but in this case the surface is too small to absorb the gases as quickly as they are liberated, and bubbles are formed, some of the energy being thus irrecoverable. If, however, the surface of the electrodes is increased by depositing platinum black upon them, the minimum electromotive force required to produce a continuous current has been found by Le Blanc [1] to be 1·67 volts.

If, instead of water, a substance such as copper sulphate had been decomposed, copper electrodes being used, we have seen (p. 64) that the amount of copper sulphate in the solution is unchanged, and in this case we find that, however small the electromotive force may be, the current is proportional to it, that is, there is no minimum electromotive force required to produce electrolysis. In any case in which the nature of the electrode is unchanged by the deposition, there is no polarisation and no back electromotive force.

It appears then that there must be at least a few free ions in the solution, since a small but limited current flows, however small the potential difference between the electrodes. According to the experiments of Kohlrausch, who investigated the relation between the electromotive force and the current in electrolytes very thoroughly, we find that any excess of electromotive force over that necessary to balance the back electromotive force due to polarisation, produces a current strictly proportional to this excess, and hence, that Ohm's law is applicable to the conduction in electrolytes. Thus $I=\frac{E-E_1}{R}$ where E is the applied E.M.F., and E_1 the back E.M.F. due to polarisation.

Electrolytic Dissociation.—It is to Arrhenius [2] that we owe a satisfactory account of the process of conduction in electrolytes. According to him, the current is entirely due to the motion of the ions in the electric field between the electrodes, and it follows that the conductivity of a solution is proportional to the number of free ions present. Thus, in the case of a solution of silver nitrate, the salt, on being dissolved, dissociates to a certain extent, and free silver ions carrying a positive charge ($Ag+$) and

[1] M. Le Blanc, *Zeitschr. phys. Chem.*, **8**, p. 299. 1891.
[2] S. Arrhenius, *Zeitschr. phys. Chem.*, **1**, p. 631. 1887.

NO_3 ions having a negative charge (NO_3-) are formed by the splitting up of the $AgNO_3$ molecules.

It is found that the conductivity of a solution diminishes on diluting it, as would be expected if the conductivity is due to the dissolved substance, for if we imagine the solution to be diluted until a given amount of dissolved substance occupies twice the original volume of solution, there will be only half the number of ions between two fixed electrodes, that is, there will only be half the number of carriers of electricity. Provided that the velocity of the ions in constant electrical field is unchanged by the act of dilution, the same electromotive force will now produce only half the transfer of electric charge in a given time, that is, half the current, so that the conductivity is now only half its value previous to dilution. Thus, if no fresh ions are produced by dilution, we should expect the conductivity to be inversely proportional to the dilution, or directly proportional to the concentration, of the dissolved substance, the concentration being for convenience taken as the number of gramme-molecules in one litre of solution.

Measurement shows, however, that the decrease in conductivity on dilution is not so great as the above simple argument would indicate; that is, the conductivity after dilution is greater than the simple proportion would give, and it therefore seems probable that, on increasing the dilution, new ions are produced by the dissociation of previously neutral molecules.

In order to follow this process, it is convenient to consider the change in the quantity $\dfrac{\textit{conductivity}}{\textit{concentration}}$ for a given solution. This new quantity is called the *Equivalent Conductivity* of the solution, and it is constant so long as the degree of dissociation is unchanged. As the solution is made more dilute, the equivalent conductivity of most of the solutions of inorganic salts in water increases, but the increase does not go on indefinitely, since a condition will eventually be reached in which all the molecules are dissociated, and hence the equivalent conductivity tends towards a superior limit for infinite dilution.

According then to Arrhenius' theory of electrolytic dissociation, the conductivity of a solution is proportional to the concentration of the free ions, and is therefore a measure of the degree of dissociation, γ, the ratio of the number of dissociated molecules to the total number. Thus, if λ_c be the equivalent conductivity of a solution at concentration c, and λ_∞ that at zero concentration, that is, infinite dilution,

$$\gamma=\frac{\lambda_c}{\lambda_\infty}.$$

For most substances before infinite dilution i by direct measuremer strongly dissociated si solution, such as the in it may be obtained w. error by extrapolation a Fig. 175 for potassium si may, however, be obtain strongly dissociated subst a knowledge of the partial as we shall see on p. 183.

The following table of mole ture 18° C. is taken from Arrh

c in gramme-equivalents per litre.	Dilution $=\frac{1}{c}$.			
	∞			
0·0001	10000			
0·0002	5000			
0·0005	2000			
0·001	1000	1		
0·002	500	1		
0·005	200	10.		
0·01	100	101		
0·02	50	99		
0·05	20	95·.		
0·1	10	92·0		
0·2	5	87·7	107·96	88·9
0·5	2	80·94	102·41	78·7
1·0	1	74·35	98·27	71·8

This theory of electrolytic dissoci ion presents many difficulties, as, for example, the presence f free ions, such as the sodium ions, in a water solution of so m chloride, since it is a well-known fact that metallic sodium a water cannot exist in contact without chemical action taking ace. But it must be remembered that an ion, that is an at with its associated charge, is in an entirely different conditio to the atom without the charge, and that if sodium be liberate by electrolysis, it is dissolved by the water as soon as it has ven up its positive charge to the cathode. There is no necessity for the ions to be imagined to be isolated in the solution; in fact it is extremely likely that they are surrounded by a number of neutral molecules of the solvent, which group is dragged along by the force on the

[1] Svante Arrhenius, "Lehrbuch der Elektro-chemie."

:t must also be remem-
ers fundamentally from
temperatures, in which
splitting up of the more
ıay be illustrated in the
ociates at high tempera-
d.

ıCl,

ed or neutral molecules.
.hloride dissociates electro-

+Cl⁻.

ı in opposition to the theory
e ions have the same velocity,
ld diffuse at greater rapidity,
of electricity, and a separation
This objection has been met
t to explain the electromotive
.olutions of the same material

e.—There is, however, ample
ıl sources to support the theory
ı't Hoff, making use of the dis-
ıce in solution exerts a pressure
.n, and that this pressure is pro-
:oncentration of the solute, showed
that this osmotic is also proportional to the absolute temperature, and theref beys the same laws as a perfect gas. From this it follows that maximum vapour pressure of water vapour over a solution[1] less than over pure water by an amount proportional to th number of molecules per litre present in the solution, and hen there is a lowering of the freezing point and a raising of th boiling-point, also proportional to the concentration of mole es of the solute. By measuring the lowering of the freezir point and raising of the boiling point, on the addition of a own mass of the solute, the molecular weight may be deter ined. In many cases the result is in accordance with that of the ordinary methods, as, for example, in the case of sugar a d similar organic substances, but for those substances which i solution form electrolytes, the pressure appears to be too great, and the molecular weight therefore too small. At great dilution, the molecular weight has, in the class of substances which dissociate into two ions, half the ordinary value, which makes it seem that there are twice the expected

[1] J. H. van't Hoff, *Zeitschr. phys. Chem.* 1, p. 481. 1887.

number of molecules present, and it is not unreasonable to suppose that in these cases the substance is completely dissociated.

When dissociation is not complete, let γ be the degree of dissociation, and n the number of ions produced by the dissociation of one molecule. Then for one gramme-molecule present per litre we have $n\gamma$ gm.-eq. of dissociated ions, and $1-\gamma$ of undissociated molecules. The concentration, counting all together, is

$$1-\gamma+n\gamma=1+(n-1)\gamma=i.$$

This is the quantity which may be determined from the freezing or boiling point experiment, and in any given case, knowing n, γ may be found. For a number of substances, the value of γ found in this way agrees very well with that found from electrolytic determinations $\left(\gamma=\frac{\lambda_c}{\lambda_\infty}\right)$, which is very strong evidence in favour of the theory of electrolytic dissociation.

In the last three columns of the following table, the values of i in the above equation are given,[1] from the three methods of observation :—

	Concentration.	i lowering of freezing point.	i from osmotic pressure.	i from electrical conductivity.
KCl . . .	0·14	—	1·81	1·86
$LiCl$ . . .	0·13	1·94	1·92	1·84
$Ca(NO_3)_2$.	0·18	2·47	2·48	2·46
$MgCl_2$ . .	0·19	2·68	2·79	2·48
$CaCl_2$. . .	0·184	2·67	2·78	2·42

Migration of the Ions.—According to the above theory, the free ions in solution experience forces due to the electric field in which they are situated, and hence acquire a velocity, the positive ions moving towards the cathode and the negative ions towards the anode. Except on the first application of the field, the ions will not have an acceleration, since in their motion they will encounter so many neutral molecules that their velocity will soon reach a limit ; just as very small falling bodies soon reach a limiting velocity owing to the viscous resistance of the air. In the case of the ions, the limiting velocity depends in the first place upon the intensity of the electric field and the charge upon the ion, but it also depends upon the nature of the solvent and upon the size of the ion with its accompanying group of neutral molecules. Since this last varies for different ions, we should expect that their velocities in equal electric fields would be

[1] J. H. van't Hoff and L. Th. Reicher, *Zeitschr. phys. Chem.*, 8, p. 198. 1889.

different, and Hittorf explained the variation in concentration of the solute at the anode and cathode which usually occurs, in terms of this difference in the velocity of the positive and negative ions, and even succeeded in determining the ratios of the velocities of migration of the two ions in a number of cases.[1]

This variation in concentration may easily be observed in the case of the electrolysis of a solution of copper sulphate using copper electrodes, the colour of the solution becoming lighter near the cathode, since the SO_4^- ions have a greater velocity than the Cu^+ ions. To observe the effect it is advantageous to use horizontal electrodes one above the other, the upper one being the cathode, with which arrangement the phenomenon is not masked by convection currents set up by the variations in density in the different parts of the cell.

The diagram given by Hittorf is a very convenient one for explaining the effect of the migration of the ions upon the changes in concentration occurring in an electrolyte. Let the dots represent positive and the circles negative ions (Fig. 176). At the instant of application of the electric field, the uniform arrangement of the two sets is indicated by the first row. Then, if the positive ions be imagined to have, for simplicity, twice the velocity of the negative ions, the state of affairs an instant later will be represented by the middle row. The deposition at the cathode is 3 ions and at the anode 3, and 7 molecules remain in solution, but of these 4 are in the cathode half of the cell and 3 in the anode half. The third row represents the cell still another instant later, and it will be seen that the total deposition at each electrode is now represented by 6 ions and that 4 molecules remain, 3 in the cathode half and 1 in the anode half. Thus at each step it will be seen that the loss in concentration of solute on the anode side of the median line is twice as great as that on the cathode side. In this simple case we can see that—

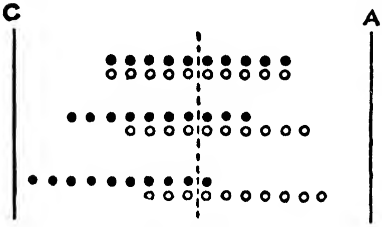


FIG. 176.

$$\frac{\text{Diminution in concentration at anode}}{\text{Diminution in concentration at cathode}} = \frac{\text{velocity of + ions}}{\text{velocity of − ions}}$$

Or in general, if u be the velocity of the positive ions and v that of the negative ions, the current and therefore the total deposition in a given time are proportional to $(u+v)$. Let the current flow for such a time that $(u+v)$ gramme-molecules of

[1] W. Hittorf, *Pogg. Ann.*, **89, 98, 103, 106.** 1853–1859.

solute are removed from the solution. Now, considering the space near the cathode $(u+v)$ gramme-equivalents of positive ions have been removed by deposition and u gained by migration, leaving a loss of

$$(u+v)-u=v.$$

Also v gramme-equivalents of negative ions are lost by migration, which shows that this part of the solution is uncharged, as it should be, and since it has lost v gramme-equivalents of both kinds of ions it has lost v gramme-molecules of the solute. Similarly, on the anode side, total loss of negative ions by deposition is $(u+v)$ gramme-equivalents, and gain by migration is v, leaving a balance of u gramme-equivalents lost. Also, loss in positive ions by migration is u gramme-equivalents, and therefore resulting loss is u gramme-molecules of solute.

Hence we obtain Hittorf's relation—

$$\frac{\text{Loss in concentration at cathode}}{\text{Loss in concentration at anode}}=\frac{v}{u}.$$

To determine experimentally the ratio $\frac{v}{u}$, or the quantity $\frac{v}{u+v}$, which is the *transport ratio* or *migration constant* of Hittorf, changes in concentration of the solute in the neighbourhood of the cathode and anode may be found by chemical means. Two beakers (Fig. 177) contain the solution, whose concentration at the start is uniform. They are connected electrically by a small syphon containing the electrolyte and through which the diffusion of the solute tending to equalise the concentrations in the two vessels will take place so slowly that a determination by chemical analysis of the amount of solvent removed from each vessel for the passage of a given current may be used to measure $\frac{v}{u}$.

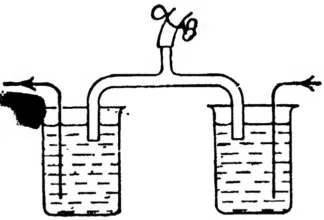


FIG. 177.

By this and similar methods, which must be suitably modified when one of the ions is soluble in water, the values of $\frac{v}{u+v}$ given in the table on p. 185 have been found.

Ionic Velocities.—A further step, due to Kohlrausch,[1] enables us to determine the sum of the actual velocities u and v in terms of the conductivity of the solution and the concentration of the

[1] F. Kohlrausch, *Wied. Ann.*, 6, p. 145. 1879.

solute. Let us take a case in which there are m gramme-equivalents of completely dissociated molecules per cubic centimetre of the solution; then the concentration of both positive and negative ions is m gramme-equivalents per cubic centimetre. Now each gramme-equivalent of positive ions carries 96,470 coulombs, and if the velocity of these ions is u, the charge passing unit cross-section of the cell in one second is $mu \,.\, 96{,}470$.

Therefore, current density due to movement of positive ions is $mu \,.\, 96{,}470$ amperes per square centimetre. Similarly, the stream of negative ions in the opposite direction constitutes a current density of $mv \,.\, 96{,}470$ amperes per square centimetre. But the effective current is the sum of these two, since they are opposite charges moving in opposite directions.

$\therefore$ resultant current density$=m(u+v)96{,}470$ amperes per square centimetre.

The same quantity may also be expressed in terms of the conductivity of the solution and the potential gradient in it. Thus the conductivity k is the inverse of the resistivity, and is the current produced in a conductor of unit cross-section and unit length, for unit potential difference between its ends; that is, it is the current density for unit potential gradient; therefore for potential gradient E volts per centimetre length

Current density$=k$E amperes per square centimetre,

$$\therefore\; m(u+v)96{,}470 = kE,$$

$$u+v=\frac{k}{m}\cdot\frac{E}{96{,}470}.$$

If now we take c to be the concentration in gramme-equivalents per litre,

$$c=1000m$$

and,
$$u+v=0{\cdot}01036\,.\,\frac{k}{c}\,.\,E.$$

The transport ratio $\frac{v}{u+v}$ being known from Hittorf's method, u and v may be separately calculated from the two equations. Further, $\frac{k}{c}$ is the quantity we have called the equivalent conductivity (p. 174) at infinite dilution, λ_∞, since we have obtained our relation on the assumption that dissociation is complete. Since this quantity is known from Kohlrausch's measurements of the conductivity of highly dissociated acids and salts, u and v for unit potential gradient are known. It will be seen from the table that these velocities are very small. They must not be confused with the velocity of the free ion in the solution, on

account of which it exerts a pressure called the osmotic pressure on the boundary of the solution, which in the case of hydrogen ions is about $18{\cdot}4 \times 10^4$ cm. per sec., the different individual ions moving indiscriminately in all directions. The velocity u is a drift of the ions towards the cathode, due to the applied electric field.

The ionic velocities increase with rising temperature.

FOR SOLUTIONS IN WATER AT 18° C.

	Partial[1] conductivity in practical C.G.S. units.	Ionic velocity[2] in cm. per sec. for potential gradient of 1 volt per cm.		Partial[1] conductivity in practical C.G.S. units.	Ionic velocity[2] in cm. per sec. for potential gradient of 1 volt per cm.
Li . .	33·4	0·000347	F . .	46·6	
Na . .	43·6	0·000451	Cl . .	65·4	0·000678
K . .	64·7	0·000670	Br . .	67·6	
Rb . .	68		I . .	66·4	0·000685
Cs . .	68		NO_3 . .	61·8	0·000640
NH_4 . .	64	0·000660	ClO_3 .	55·0	
Ag . .	54·0	0·000570	IO_3 . .	33·9	
½Zn . .	46·7		BrO_3 .	46	
½Mg . .	46·0		ClO_4 . .	64	
½Ba . .	55·5		IO_4 . .	48	
½Pb . .	61·3		½CrO_4 .	72	
H . .	318	0·003250	½SO_4 . .	68·4	
½Sr . .	51·7		OH . .	174	0·001780
½Cu . .	47·3				
½Ca . .	51·8				

Direct Determination of Ionic Velocity.—Several direct determinations of ionic velocities have been made, the general method of which is to follow the course of an ion by means of some chemical reaction produced by it. The results are in fair agreement with the determinations of Kohlrausch.

Sir Oliver Lodge[3] filled two vessels with an electrolyte, and joined the two by a horizontal tube containing a solution of some suitable material in gelatine or in solid agar-agar jelly. Using a weak solution of sulphuric acid in the vessels, and sodium chloride with phenolphthalein as an indicator in the tube: on passing the current from one vessel to the other the H^+ ions form HCl with the sodium chloride, and decolourise the phenolphthalein. The progress of the H^+ ions could thus be watched, and their velocity measured. In another experiment, using $BaCl_2$ solution in the vessels and acetic acid and silver sulphate in the gelatine, the progress of the Ba^+ ions could be observed by the precipitate of $BaSO_4$, and of the Cl^- by the precipitate of AgCl.

[1] F. Kohlrausch, "Lehrbuch der Praktischen Physik."
[2] Svante Arrhenius, "Lehrbuch der Elektro-chemie."
[3] O. Lodge, Brit. Assoc., Birmingham, 1886.

W. C. D. Whetham [1] used two solutions differing in density, but of the same conductivity and having one ion in common. Thus with deci-normal solutions of potassium bichromate and potassium carbonate, the K^+ ions pass in one direction, and the $Cr_2O_7^-$ and CO_3^- in the opposite direction. Since the colour of the bichromate is due to the $Cr_2O_7^-$ ions, the travel of the surface of separation of the two liquids can be observed.

B. D. Steele [2] has further modified the method by avoiding the use of colouring matter, the surfaces of separation of the liquids being sufficiently well defined on account of their different refractive indices, due to the slight differences in density produced on replacing one ion by another. The applicability of the method is thus considerably extended since it is not necessary to depend upon a coloured indicator, of which there are only a few that are suitable. The salt solution under examination is placed in a U-tube, and is bounded at the two ends by a gelatine solution containing the indicators employed. In some of the experiments lithium chloride and sodium acetate are used, the former at the anode and the latter at the cathode. Using, for example, potassium chloride as the salt in solution the lithium and potassium ions travel in the direction of the current, the potassium chloride being converted to lithium chloride. With the ion in solution having slightly higher velocity than the indicating ion that follows it, a very clear surface of separation can be observed with suitable illumination, and its velocity is that of the more rapidly moving ion, in this case potassium. In a similar manner, at the other end of the column of solution the anions travel from the gelatine, and the chloride is converted into acetate, and the velocity of travel of the surface of separation is that of the chlorine ions. The arrangement is always such that the denser liquid lies underneath the less dense, so that the surfaces of separation are not disturbed by convection currents, and when necessary for this, the U-tube is of the inverted form. Precautions are taken that the specific resistance of the solution shall be as nearly uniform as possible, as only then is the potential gradient throughout the solution known.

Partial or Ionic Conductivities.—On examining the equation

$$u+v=0{\cdot}01036\frac{k}{c}\,.\,E, \qquad \text{(p. 180)}$$

we see that for any given value of E, the quantity $\frac{k}{c}$ or λ_∞, which may be written $\frac{u+v}{0{\cdot}01036E}$, is the sum of two others, $\frac{u}{0{\cdot}01036E}$ and

[1] W. C. D. Whetham, *Proc. Roy. Soc.*, 52, p. 283 (1892); 58, p. 182 (1895).
[2] B. D. Steele, *Chem. Soc. Journ.*. 79. p. 414. 1901.

$\dfrac{v}{0{\cdot}01036\text{E}}$, which are called the *partial or ionic conductivities* of the two ions. In any case λ_∞ is made up of the sum of two partial conductivities whose ratio is $u : v$, and hence if λ_∞ can be measured and also the transport ratios, the partial conductivities can be found. *The partial conductivity of any ion is independent of the other ions in the solution,* and hence, if the partial conductivities of the ions in a solution are known, the equivalent conductivity at infinite dilution λ_∞ is known, since it is the sum of the partial conductivities.

Many partial conductivities are given in the table on p. 181, and from them, the limiting equivalent conductivity λ_∞ of a substance which is only partially dissociated at very great dilution may be found.

Measurement of Conductivity.—The difficulty met with in measuring the conductivity of electrolytes is due to the polarisation which generally occurs, producing a back electromotive force that cannot always be separated from the ohmic potential difference corresponding to the resistance of the electrolyte. In some cases the electromotive force due to polarisation may be eliminated by using electrodes of the material which is present in the ionic state in the solution. Thus in the case of a solution of copper sulphate, copper electrodes may be used, and a method of simple substitution, due to Horsford, employed. The tube containing the solution is placed vertically, and is provided with disc electrodes, which nearly fill the cross-section of the tube (Fig. 178). The resistance in the box R is adjusted until the galvanometer deflection is a convenient amount. The upper electrode is then pushed downwards by a measured distance l, and R is adjusted to give the same deflection as before. The resistance of the length l of the electrolyte in the tube is equal to the change of resistance in R. The mean area of cross-section of the tube may be determined by finding the weight of water required to fill a measured length of it.

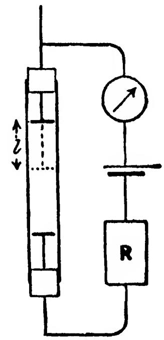


FIG. 178.

The method most generally applied is due to Kohlrausch. The electrolytic cell, which has the form shown in Fig. 179 for good conductors, and Fig. 180 for bad conductors, is placed in one arm of a slide-wire bridge, a known resistance being placed in the other. The polarisation electromotive force may be greatly reduced by using electrodes of large area, since for a given amount of deposition the layer of deposit is then thinner than when a small electrode is used. The effective area is much increased, in the case of platinum electrodes, on covering them

with a layer of platinum black, by immersing them in a solution of platinum chloride, and passing a current backwards and forwards through the solution a number of times.

In order to reduce the polarisation still further, a small and rapidly alternating current is used, so that the small amount of deposition occurring when the current passes in one direction will be removed on its reversal. A small induction coil with a high frequency trembler is a very efficient source of electromotive force, but in this case an ordinary galvanometer is useless for finding the position of balance, since the deflection is proportional to the first power of the current and would be reversed with it;

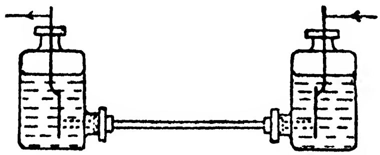


Fig. 179.

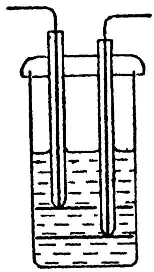


Fig. 180.

hence, there will be no deflection with an alternating current. In order to get over this difficulty, a telephone receiver is used instead of a galvanometer, and the observer adjusts the position of the slide-wire contact until a minimum of sound is heard in the telephone. Alternating current galvanometers, such as the Duddell thermo-galvanometer described on p. 75, have also been used, which instrument is capable of detecting very small alternating currents.

If resistance coils are used as standards, they should be few in number and should be wound so that they have as small an inductance and capacity as possible, as, otherwise, there will not be a perfect balance when the proportionality in resistance of the four resistances of the Wheatstone's bridge is attained. The higher the frequency of the alternating current the greater will be the disturbance due to this cause. With frequencies below 200 alternations per second the disturbance is inappreciable when ordinary resistance coils are used.

In using a cell of the type shown in Fig. 179, the cross-section of the tube may be found by means of a mercury thread whose length in the tube and whose mass are measured. The indeterminate resistance where the end of the narrow tube enters the vessel may be eliminated by performing the experiment twice,

using two different lengths of tube, cut from the same piece. As the end errors are the same for each tube, the difference in the two resistances found is equal to that of a column of length equal to the difference in length of the tubes.

If the cell have the form shown in Fig. 180, the absolute conductivity cannot be found from the dimensions of the liquid between the electrodes, with any degree of accuracy. It is usual then to find the resistance first with a standard electrolyte of known conductivity, and then with that whose conductivity it is required to find. For this purpose Kohlrausch[1] gives the conductivities shown in the following table :—

NORMAL SOLUTIONS IN WATER AT 18° C.

	$k=\frac{1}{S}$. (Ohms and cm.s.)	$\frac{1}{k}\cdot\frac{dk}{dt}$.	$\frac{v}{u+v}$.
KOH	1840×10^{-4}	0·0186	0·74
KCl	$982{\cdot}6 \times 10^{-4}$	0·0193	0·51
KBr	1030×10^{-4}	0·0190	0·51
KI	1036×10^{-4}	0·0190	0·51
KNO_3 . . .	805×10^{-4}	0·0200	0·49
$\frac{1}{2}K_2SO_4$ . . .	$715{\cdot}9 \times 10^{-4}$	0·0205	0·50
NH_4Cl . . .	970×10^{-4}	0·0194	0·51
NaOH . . .	1600×10^{-4}	0·0197	0·83
NaCl	$743{\cdot}5 \times 10^{-4}$	0·0212	0·64
$NaNO_3$ . . .	659×10^{-4}	0·0215	0·61
$\frac{1}{2}Na_2SO_4$ . . .	508×10^{-4}	0·0236	0·64
$\frac{1}{2}ZnCl_2$ . . .	550×10^{-4}	0·0220	0·70
$\frac{1}{2}ZnSO_4$ . . .	$262{\cdot}1 \times 10^{-4}$	0·0218	0·68
$\frac{1}{2}CuSO_4$ . . .	$257{\cdot}7 \times 10^{-4}$	0·0216	0·70
$AgNO_3$ . . .	676×10^{-4}	0·0210	0·50
HCl	3000×10^{-4}	0·0159	0·17
HNO_3 . . .	2990×10^{-4}	0·0150	0·17
$\frac{1}{2}H_2SO_4$ . . .	1970×10^{-4}	0·0120	0·17

	$k=\frac{1}{S}$. (Ohms and cm.s.)	$\frac{1}{k}\cdot\frac{dk}{dt}$.		$k=\frac{1}{S}$.	$\frac{1}{k}\cdot\frac{dk}{dt}$.
KCl 5 per cent.	690×10^{-4}	0·020	K_2SO_4 5 per cent.	460×10^{-4}	0·022
" 10 "	1360×10^{-4}	0·019	" 10 "	860×10^{-4}	0·020
" 15 "	2020×10^{-4}	0·018	$ZnSO_4$ 5 "	191×10^{-4}	0·022
" 20 "	2680×10^{-4}	0·017	" 10 "	321×10^{-4}	0·022
$CuSO_4$ 5 "	189×10^{-4}	0·022	" 15 "	415×10^{-4}	0·023
" 10 "	320×10^{-4}	0·022	" 20 "	470×10^{-4}	0·024
" 15 "	421×10^{-4}	0·023	" 25 "	480×10^{-4}	0·026
			" 30 "	440×10^{-4}	0·027

[1] F. Kohlrausch, "Lehrbuch der Praktischen Physik."

In all cases it is necessary to observe the temperature of the electrolyte at the time of measurement, since the resistance falls about 2·4 per cent. for a rise in temperature of one degree when the temperature is 18° C.

Application of Thermodynamics to Reversible Cells.—The second law of thermodynamics can only be applied to processes which are strictly reversible, that is to say, will proceed in either direction when one of the forces producing equilibrium is increased by an indefinitely small amount. A gas enclosed in a cylinder by means of a frictionless piston affords a good example, for if the pressure inside the cylinder exceed that outside by ever so small an amount, the piston is driven outwards, and when the pressure outside is greater than that inside by however small an amount, the piston moves inwards. If the piston is not frictionless, it requires a finite difference of pressure on the two sides to move it, and the work done in moving it in opposition to the force of friction is irrecoverable; the process is then irreversible in the thermodynamic sense.

In the case of an electric cell, we have reversibility when there is no polarisation, as in the case of a Daniell's cell, since at each electrode the ion liberated does not alter the chemical nature of the electrode, and we have already seen (p. 173) that in such a case an electromotive force, however small, will produce a current. It is also evident that if the current be allowed to flow until a certain amount of anode is dissolved and an equivalent amount of metal is deposited on the cathode, we can, on reversing the current by some external means, bring the cell back again to its original condition. This in itself is a satisfactory test for reversibility in the case of a cell.

Whenever a current flows, an amount of work i^2rt is converted into heat in the cell, and this is irreversible, since the process cannot be reversed; that is, the application of heat will not produce the current. Another source of irreversibility is the diffusion that takes place when there are two liquids in the cell, but we shall assume the currents and times taken to be small enough to justify us in neglecting these two irreversible processes.

Let us then consider a reversible cell whose electromotive force is e at absolute temperature T, which can be maintained constant, and let the cell produce current until a charge q has passed round the circuit. Drawing an indicator diagram for the process (Fig. 181), PQ represents the passage of the charge q round the circuit, and this line is parallel to the axis of q, since the electromotive force is constant at constant temperature. Now thermally isolate the cell, and let a further infinitesimal charge pass; the only source of energy is now the cell itself, and let us suppose that the using up of the energy of the cell causes drop of tempera-

ture δT. The temperature is now $T-\delta T$, and the electromotive force $e-\frac{de}{dT}\delta T$, where $\frac{de}{dT}$ is the rate of change of electromotive force with temperature. On our diagram this change is represented by the path QR. Now, maintaining the lower temperature constant, pass a current in the opposite direction to the first, until the charge q has passed through the cell; this brings us to the point S. Then pass a sufficient charge to bring the cell, when isolated thermally from outside sources, back to its original temperature T.

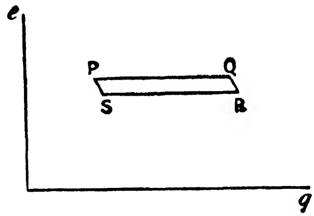


FIG. 181.

If all the processes are carried out by indefinitely small differences between the electromotive force of the cell and the applied external electromotive force, every part of the cycle is reversible, and the two adiabatic processes represented by QR and SP are identical and the cycle is complete.

It is shown in works on thermodynamics that in any reversible cycle between two temperatures, the ratio of the useful work performed during the cycle to the heat drawn from the source at the higher temperature, is equal to the ratio of the difference in the two temperatures to that of the source, or

$$\frac{h-h_1}{h}=\frac{T-T_1}{T}.$$

The work done by the cell during the process PQ is eq, and that restored to the cell during the process RS is $\left(e-\frac{de}{dT}\delta T\right)q$, and if δT is so small that the difference in the amounts of work represented by the processes QR and SP is infinitesimal, the balance of useful work done by the cell is $eq-\left(e-\frac{de}{dT}\delta T\right)q=q\,.\,\delta T\,.\,\frac{de}{dT}$, and this is equal to $h-h_1$, the excess of heat absorbed over that given up, so that we have from thermodynamics,

$$\frac{q\,.\,\delta T\,.\,\frac{de}{dT}}{h}=\frac{\delta T}{T},$$

from which,

$$h=qT\frac{de}{dT}.$$

This relation holds, whatever the chemical changes going on in the cell, since, being reversible, it is brought back to its original

condition on completing the cycle, for as much charge has passed through it in one direction as in the other.

The actual heat h drawn from the source depends upon the work done, eq, and the energy supplied by the chemical reactions in the cell. If H be the amount of heat measured in ergs, which is liberated by the chemical processes occurring when a unit of charge passes through the cell, Hq is the amount liberated during the process PQ, and the work done being eq, we have by the principle of the conservation of energy

$$eq = \mathrm{H}q + h,$$

that is,

$$h = eq - \mathrm{H}q,$$

and substituting this value of h in our previous equation we get

$$eq - \mathrm{H}q = q\mathrm{T}\frac{de}{d\mathrm{T}},$$

or,

$$e = \mathrm{H} + \mathrm{T}\frac{de}{d\mathrm{T}}.$$

This is known as the equation of Helmholtz. From it we see that when the temperature coefficient $\frac{de}{d\mathrm{T}}$ is zero, $e = \mathrm{H}$, and the energy of the current is exactly supplied by the chemical reactions occurring in the cell. This is approximately the case in the Daniell's cell, in which case $\mathrm{H} = 2{\cdot}66 \times 4{\cdot}18 \times 10^7 = 1{\cdot}112 \times 10^8$ C.G.S. units, and therefore $e = 1{\cdot}112$ volts. 2·66 is the number of calories liberated when one equivalent of zinc (0·00338 grammes) replaces an equivalent amount (0·00329) of copper in the sulphate. The observed electromotive force of the Daniell is about 1·09 volts.

If the electromotive force of the cell increases with rise in temperature, $\frac{de}{d\mathrm{T}}$ is positive and $e > \mathrm{H}$. Hence, in order to supply the energy necessary to maintain the current, the heat of the cell itself is drawn upon, and the cell is thereby cooled. On the other hand, if the electromotive force falls with rising temperature, $\frac{de}{d\mathrm{T}}$ is negative and $e < \mathrm{H}$. In this case the energy liberated by the chemical reaction is greater than that required by the current and the cell gets warmer when running.

Jahn [1] determined experimentally the electromotive force of a number of cells, and their temperature coefficients at a number of temperatures, also the heats of chemical reaction by means of the ice calorimeter, and found the results to be in accordance with the equation of Helmholtz.

Standard Cells.—The two most important cells used as stan-

[1] H. Jahn, *Wied. Ann.*, **28**, p. 491. 1886.

dards of electromotive force are of the reversible type, thus ensuring constancy of electromotive force and temperature coefficient; they are the Latimer-Clark cell and the Weston or Cadmium cell. There are many patterns of these cells; one very useful pattern of the Clark cell due to Lord Rayleigh is shown in Fig. 182. Through the bottom of each limb of the H-shaped tube is sealed a platinum wire to serve as terminal. Mercury is poured into one limb, and upon this rests a paste consisting of mercurous sulphate and zinc sulphate, and in the other is an amalgam of zinc (10 per cent. Zn), upon which rests a layer of crystals of zinc sulphate. Zinc sulphate solution fills the tubes above the cross-piece, and the whole is sealed up with corks and paraffin wax.

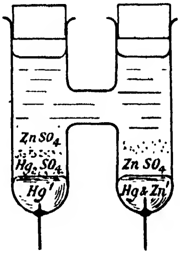


FIG. 182.

According to Jäger and Kahle (Reichsanstalt), the E.M.F. is

$$1{\cdot}4328-0{\cdot}00119(t-15)-0{\cdot}000007(t-15)^2 \text{ volt,}$$

where t is the temperature Centigrade.

In the Weston or Cadmium cell, cadmium amalgam and sulphate replace the zinc amalgam and sulphate of the Clark cell. In 1908, the International Conference on Electrical Units and Standards adopted 1·0184 volt as the electromotive force of a cell of this type at 20° C., the amalgam containing from 10 to 13 per cent. of cadmium. This value was subsequently revised to 1·0183 volt and this figure fixes the "international volt," although it was the original intention that the unit of potential difference should be derived from those of current and resistance. The international units served a very useful purpose for many years, but reference is now (since 1948, by international agreement) made to the absolute units. Since the standard Weston cell is found to have at 20° C. an electromotive force of 1·01864 absolute volt, the international volt is 1·01864/1·01830=1·00034 absolute volt.

The Weston cell replaced the Clark cell because it lasts better. In addition, it has a smaller temperature-coefficient. In fact, the E.M.F. at t° C. exceeds that at 20° C. by

$$-39{\cdot}39\times10^{-6}(t-20)-0{\cdot}903\times10^{-6}(t-20)^2$$
$$+0{\cdot}0066\times10^{-6}(t-20)^3-0{\cdot}00015\times10^{-6}(t-20)^4 \text{ volt.[1]}$$

Standard cells must not deliver currents above a microampere or so.

Concentration Cells.—The possibility of constructing cells in which the source of energy is not due to chemical action but to the diffusion occurring between two solutions of the same sub-

[1] For details, see P. Vigoureux and C. E. Webb, "Principles of Electric and Magnetic Measurements," p. 14. 1947.

stance at different concentrations was first pointed out by Helmholtz.[1] In both the cells indicated by the formulæ

$$\mathrm{Cu} \mid \mathrm{CuSO_4}\ \text{(concentrated)} \mid \mathrm{CuSO_4}\ \text{(dilute)} \mid \mathrm{Cu}$$

and, $$\mathrm{Ag} \mid \mathrm{AgNO_3}\ \text{(concentrated)} \mid \mathrm{AgNO_3}\ \text{(dilute)} \mid \mathrm{Ag}$$

the metal in contact with the dilute solution goes into solution, and that in contact with the concentrated solution receives a deposit when the cell is in action. As the metallic ions are positive the latter becomes the positive pole of the cell.

In order to examine the mode of change of the concentration of the two solutions, let us consider that $u+v$ gramme-equivalent of metal is dissolved at the anode, and an equal quantity deposited at the cathode, where the ionic velocities of the positive and negative ions are respectively u and v. We have for the silver cell, as on p. 179—

Loss of Ag at cathode by deposition $=u+v$,
Gain ,, ,, transport $=u$,
$\therefore$ total loss of Ag $=(u+v)-u=v$.
Also loss of NO_3 by transport $=v$,
$\therefore$ loss of $AgNO_3=v$ gramme-molecules.

On the other hand, at the anode we have—

Gain of Ag by solution $=u+v$,
Loss of Ag by transport $=u$,
$\therefore$ gain of Ag $=(u+v)-u=v$.

And since there is a gain of NO_3 by transport equal to this—

Gain in $AgNO_3$ at anode $=v$ gramme-molecules.

Thus the result of the process is a transference of v gramme-molecules of $AgNO_3$ from the concentrated to the dilute solution. If instead of $u+v$ gramme-atoms deposited we take one gramme-atom, the transference of $AgNO_3$ is equal to $\frac{v}{u+v}$ gramme-molecules, and $\frac{v}{u+v}$ is the transport ratio of the negative ion. Other cells have been devised in which the migration of the positive ion has been employed, and in this case the transference of the salt for one gramme-equivalent of deposit would be $\frac{u}{u+v}$ gramme-molecules. Putting the transport ratio $\frac{v}{u+v}$ equal to n, we see that $\frac{u}{u+v}=1-n$.

[1] H. Helmholtz, *Wied. Ann.*, **3**, p. 201. 1878.

Source of Energy in Concentration Cells.—We may seek for the source of energy of the current in the diluting of the solution from the concentration at one electrode (C_1) to the less concentration (C_2) at the other electrode. The substance in solution exerts a pressure, the osmotic pressure, which has been shown by Pfeffer, and by van't Hoff (see p. 176) to have the same value as that exerted by an equal number of molecules existing as a gas in a space equal in volume to the solution, and hence work is done as the solute expands from molecular volume $\frac{1}{C_1}$ to molecular volume $\frac{1}{C_2}$.

The pressure of a gas is given by the relation $PV=RT$, where T is the absolute temperature, and R a constant that can be found from the volume of a given amount of gas at some standard temperature and pressure.

It is usual to take V as the reciprocal of the concentration in gramme-molecules per unit volume, so that if this be known, R will enable us to calculate the osmotic pressure P at any temperature T. If then we imagine the expansion of the solute to take place reversibly by enclosing it in a cylinder, in which works a piston constructed of a medium which is permeable to the solvent but not to the solute, so that the osmotic pressure P may be balanced by an external pressure very slightly less than P, applied to the piston—

work for small increase dV in volume$=PdV$,

$$\begin{aligned}\therefore \text{ total work} &= \int_{P_1}^{P_2} PdV,\\ &= \int_{V_1}^{V_2} \frac{RT}{V}\,dV,\\ &= RT[\log_\epsilon V]_{V_1}^{V_2},\\ &= RT\log_\epsilon \frac{V_2}{V_1},\end{aligned}$$

or remembering that, $V_1=\frac{1}{C_1}$, and, $V_2=\frac{1}{C_2}$,

$$\text{work} = RT\log_\epsilon \frac{C_1}{C_2}.$$

Such semi-permeable membranes have only been found for a few substances, but the actual work done by the solute in expanding so that the solution becomes more dilute does not depend upon the mechanical method of carrying out the dilution,

provided that the process is reversible, which it is in the case of the concentration cell; for the solute may be carried back from the weak to the strong part of the solution on reversing the current by means of some external electromotive force.

E.M.F. of Concentration Cells.—Now, work performed in producing current $=eq$ ergs, where q is the charge in absolute units (9647), equivalent to the deposition of one gramme-equivalent of ion, and as there is no other source of energy than the work done by the solute in changing from concentration C_1 to concentration C_2, and remembering that when the solute is completely dissociated, the osmotic pressure is double that for no dissociation, by Avogadro's law—

$$\text{Work}=2\text{RT}\log_e\frac{C_1}{C_2}$$

and therefore, for n gramme-equivalents of solute transferred by the passage of the above charge, we may write

$$eq=2n\text{RT}\log_e\frac{C_1}{C_2}.$$

Taking the density of hydrogen as 0·0000899 gramme per cubic centimetre at 0°, and the atmospheric pressure of $76\times980{\cdot}6\times13{\cdot}59$ dynes per square centimetre, the molecular weight being 2·016, the molecular volume (V) is $\frac{2{\cdot}016}{0{\cdot}0000899}$, and since PV=RT—

$$76\times980{\cdot}6\times13{\cdot}59\times\frac{2{\cdot}016}{0{\cdot}0000899}=\text{RT},$$

from which, $\text{R}=8{\cdot}32\times10^7.$

And remembering that $\log_e\frac{C_1}{C_2}=2{\cdot}303\log_{10}\frac{C_1}{C_2}$, we have—

$$e=\frac{8{\cdot}32\times10^7\times2{\cdot}303\times2}{9647}n\text{T}\log_{10}\frac{C_1}{C_2},$$

And at 18° C., T=291,

$$\therefore\ e=5{\cdot}78\times2n\log_{10}\frac{C_1}{C_2}\times10^6,$$

$$\text{E}=0{\cdot}0578\times2n\log_{10}\frac{C_1}{C_2}\text{ volts.}$$

Thus the electromotive force of a concentration cell consisting of two solutions is proportional to the absolute temperature, and depends upon the ratio only of the concentrations. In the silver nitrate cell suggested above, $n=\frac{61{\cdot}8}{54+61{\cdot}8}=0{\cdot}533$ (see table on

p. 181), and at temperature 18° C., with ratio of concentrations 10 : 1,

$$E = 0{\cdot}0578 \times 2 \times 0{\cdot}533 = 0{\cdot}0615 \text{ volt.}$$

The value found by Nernst is 0·055 volt at 18° C., and the discrepancy between this and the calculated value he attributed to the incomplete dissociation of the salt in solution.

Since there is no resultant chemical reaction in the concentration cell, the total amount of solute remaining constant, and all the processes being reversible, we may apply Helmholtz's E.M.F. equation of p. 188—

putting $H=0$,

$$\text{then, } e = T\frac{de}{dT},$$

$$\text{or, } \frac{de}{e} = \frac{dT}{T}.$$

And integrating—

$$\log_\epsilon e = \log_\epsilon T + \text{constant},$$

$$\log_\epsilon \frac{e}{T} = \text{constant},$$

or e is proportional to T, which is in accordance with the relation found.

Amalgam Concentration Cells.—Another form of concentration cell in which the electrodes are amalgams, the two having different concentrations of the metal, and the electrolyte being a solution of some salt of the metal, was constructed by G. Meyer.[1] The electromotive force acts in such a direction that the metal is transferred from the amalgam of greater to that of less concentration, and since it has been shown that the osmotic pressure of the metal in an amalgam is proportional to the concentration, the electromotive force may be calculated as above. For cells with amalgams of Zn, Cd, Pb, Sn, Cu and Na, there was agreement between the calculated and the observed electromotive forces.

Now the expression for the work done by a gas or a solute in expanding, namely

$$\int_{P_1}^{P_2} P dV, \text{ or, } \int_{V_1}^{V_2} \frac{RT}{V} dV \qquad \text{(p. 191)}$$

depends upon the value of P, or of $\frac{RT}{V}$,

and the value of the gas constant R is calculated on the assump-

[1] G. Meyer, *Zeitschr. Phys. Chem.*, **7**, p. 477. 1891.

tion that the molecule is undissociated; but it follows that when dissociation occurs, the pressure increases in accordance with Avogadro's rule, that at equal temperatures and pressures, equal volumes contain equal numbers of molecules. Hence if each molecule dissociates into n others, on going into solution $P=n\frac{RT}{V}$.

And the expression for the work done on expansion becomes

$$\int_{V_1}^{V_2}\frac{nRT}{V}dV=nRT\log_\epsilon\frac{V_2}{V_1},$$

or since PV=constant, at constant temperature—

$$\text{work}=nRT\log_\epsilon\frac{P_1}{P_2},$$

and the value of the electromotive force at the contact of the concentrated amalgam with the solution is given by

$$eqr=nRT\log_\epsilon\frac{P_1}{P_2},$$

where qr is the quantity of electricity which passes from the concentrated amalgam to the solution for one gramme-atom of the metal to be transferred, which is $9647r$ absolute C.G.S. units; r being the valency.

$$\therefore\ e=\frac{nRT}{9647r}\log_\epsilon\frac{P_1}{P_2},$$

P_1 and P_2 being the osmotic pressures of the metal in the amalgam, and in the solution respectively.

Similarly, at the electrode of weaker amalgam—

$$e=\frac{nRT}{9647r}\log_\epsilon\frac{P_3}{P_2},$$

where P_3 is the osmotic pressure in this amalgam, and the electromotive force acts from the amalgam to the solution. The resulting electromotive force due to the whole cell is therefore the difference between these two;

i.e.
$$\frac{nRT}{9647r}\log_\epsilon\frac{P_1}{P_3}.$$

The results of Meyer's experiments agree with those calculated on the assumption that the molecular weight of the metal in the amalgam is equal to its atomic weight, the osmotic pressure being proportional to the concentration.

A further type of concentration cell may be employed, namely, one in which a gas, such as hydrogen, is dissolved, or rather

occluded by platinum or palladium. On immersion in a solution of sulphuric acid there is an electromotive force between the electrode and the solution, owing to the difference in osmotic pressure between the occluded gas in the electrode and the ions in the solution. Wulf [1] has shown that up to pressures of 1000 atmospheres, the observed and the calculated electromotive forces are in agreement.

Solution Pressure.—The consideration of the relation between osmotic pressure and electromotive force, particularly in the case of amalgams, enabled Nernst to make a great step forward in the theory of the voltaic cell. Ever since its discovery by Volta there had been a controversy regarding the location of the electromotive force in the circuit. While, on one hand, Volta and his followers maintained the junction of the metals to be the seat of the electromotive force; on the other hand, Davy looked to the contact of the metal and the solution, and explained the electromotive force in terms of the chemical affinity of the metal for the acid in solution. The experiment of the electrification of a copper and a zinc disc when placed in contact and then separated, seemed to bear out Volta's contact hypothesis, but this effect was explained by the chemical school of physicists on the ground of the difference in chemical affinity of copper and zinc for the oxygen of the air.

Following the experimental work on concentration cells, Nernst [2] explained the electromotive force in terms of the work done by the ions in travelling from places of higher to places of lower concentration, on account of the osmotic pressure exerted by them. When a metal is placed in a solution, the osmotic pressure of the ions in solution drives the ions upon the metal, but the ions in the metal itself having a certain pressure tending to drive them into solution, there will be equilibrium when these two pressures are equal.

Thus for every metal there is a particular osmotic pressure of the metallic ions in solution, for which neither deposition nor dissolving will occur. This is called the solution pressure of the metal for the given solvent. From the reasoning given above for concentration cells, it follows that the electromotive force at the contact of a metal with its solution is

$$e = \frac{RT}{r9647} \log_\epsilon \frac{P}{p} \text{ absolute units,}$$

where r is the valency of the metal, P the solution pressure, and p the osmotic pressure of the ion in solution.

The electromotive force directed from a metal to the normal

[1] T. Wulf, *Zeitschr. f. phys. Chem.*, **48**, p. 87. 1904.
[2] W. Nernst, *Zeitschr. Phys. Chem.*, **4**, p. 129. 1889.

solution of its salt has been determined by Ostwald, the values being for—

Mg, +1·22 volt.	Pb, −0·10 volt.
Zn, +0·51 ,,	H, −0·25 ,,
Al, +0·22 ,,	Cu, −0·60 ,,
Cd, +0·19 ,,	Hg, −0·99 ,,
Fe, +0·06 ,,	Ag, −1·01 ,,

From the last equation, we may find the solution pressure P for the metal, if we take p to be the osmotic pressure due to the metallic ions in the normal solution. In the case of hydrogen, taking two atoms to the molecule in the gaseous state, and the density 0·0899 gramme per litre at 0° C. under the atmospheric pressure, for a normal solution of 1 gramme per litre the pressure is $\frac{2}{0{\cdot}0899}$ atmospheres at 0° C., from van't Hoff's law (p. 176), taking the molecules to be monatomic in solution.

Hence the electromotive force between hydrogen (H_2) occluded by palladium and a solution containing 1 gramme of hydrogen per litre being $-0{\cdot}25$ volt, or $-0{\cdot}25 \times 10^8$ absolute units,

$$-0{\cdot}25 \times 10^8 = \frac{8{\cdot}32 \times 10^7 \times 273}{1 \times 9647} \log_e \frac{P}{p}$$

from which, $\log_e \frac{P}{p} = -10{\cdot}74$, and $\frac{P}{p} = 2{\cdot}16 \times 10^{-5}$

$$\therefore \; P = \frac{2 \times 2{\cdot}16 \times 10^{-5}}{0{\cdot}0899} = 4{\cdot}8 \times 10^{-4} \text{ atmosphere.}$$

The following are the approximate values of the solution pressures in atmospheres :—

Mg, 10^{23} atmospheres.	Pb, 10^{-1} atmospheres.
Zn, 10^{10} ,,	H, 10^{-5} ,,
Al, 10^{5} ,,	Cu, 10^{-10} ,,
Cd, 10^{5} ,,	Hg, 10^{-16} ,,
Fe, 10^{2} ,,	Ag, 10^{-16} ,,

The electromotive force of a reversible cell may then be represented in terms of the solution pressures of the electrodes and the osmotic pressures of the ions in solution; thus, in the case of the Daniell cell, neglecting the electromotive force due to the contact of the solutions, which is very small compared with that due to the electrodes—

$$e = \frac{RT}{2 \times 9647} \log_e \frac{P}{p} - \frac{RT}{2 \times 9647} \log_e \frac{P'}{p'}$$

where P and P′ are the solution pressures of zinc and copper, p and p' being the osmotic pressures of the zinc and copper ions in the solutions; and, taking the e.m.f. from zinc to solution as

0·51 volt, and from solution to copper to be −0·60 (table, p. 196), the E.M.F. of the cell would be 1·11 volt.

Dropping Electrodes.—At the moment of bringing a metal into contact with a solution, a transference of ions will take place in accordance with the respective osmotic and solution pressures. If the circuit is not completed through some other electrode, this transference of the ions will produce an electromotive force at the contact of the metal and the solution, which will increase as the transference of ions continues, until a limit is reached, at which further transference is prevented. Thus, in the case of mercury in contact with a solution of sulphuric acid, it will be seen from the table of solution pressures that the mercury ions in the solution will pass to the jet (Fig. 183), since the solution pressure of the mercury is extremely small, and is less than that for even a very weak solution. The mercury will then become charged positively with respect to the solution. This electromotive force will, however, require an appreciable time for its establishment, and it was pointed out by Helmholtz that if the mercury issue from a small orifice in the form of a jet into the solution, it will break up into drops before any appreciable potential difference is produced, and the jet will then be at the same potential as the solution. Although the explanation of the production of the p.d. given by Helmholtz differs from the above, the fact of the usefulness of the dropping electrode remains.

It will easily be seen that any attempt to measure the electromotive force at a single contact between a metal and an electrolyte is frustrated by the necessity of making electrical connection with the electrolyte by means of a second metallic electrode. The measurement gives us merely the algebraic sum of the two respective electromotive forces between the metals and the solution, unless we can obtain some electrode at which there is no electromotive force at its contact with the solution. The dropping electrode of Helmholtz supplies us with such a terminal, but the observations made with it were very unreliable until Paschen [1] constructed a form of it in which the mercury jet does not enter the liquid until it is just about to break up, the interval of time in which the mercury is in contact with the solution before breaking into drops being then a minimum. In Fig. 183, the leads 1 and 2 go to the mercury of the jet and the pool of mercury in the vessel. On measuring the difference of potential between 1 and 2, we thus obtain that between the mercury at rest in the pool at the bottom and the solution.

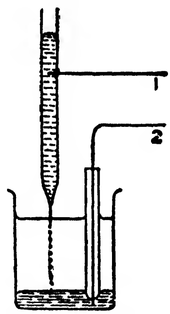


FIG. 183.

[1] F. Paschen, *Wied. Ann.* **41**, p. 42. 1890. For a recent study, see R. J. Newcombe, *Nature*, **169**, p. 240. 1952.

Once this difference of potential between the mercury and the solution has been found, the dropping electrode can be dispensed with, since the resulting difference of potential, produced by combining this with other electrodes, can be found by the ordinary means.

Normal Electrode.—It is convenient to employ some constant form of electrode of useful pattern and known electromotive force, to combine with other electrodes whose electromotive force it is required to find. This usually consists of mercury upon which rests a layer of mercurous chloride, and again upon this a normal solution of potassium chloride, which fills the remainder of the vessel and the tube T (Fig. 184). The latter can be placed in any cell in which the electromotive force at contact with the electrode and the solution is required. The resultant electromotive force being measured in the ordinary way, that of the normal electrode may be deducted and the unknown electromotive force obtained.

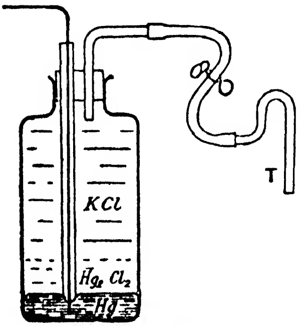


FIG. 184.

The electromotive force of the normal electrode is 0·56 volt, the mercury being positive to the solution.

Capillary Electrometers.—An alternative method of measuring the electromotive force at the contact of mercury and a dilute solution of sulphuric acid has been devised by Lippmann.[1] We have seen that the electromotive force is so directed that the mercury is positive with respect to the solution. Now it is found that the surface tension of the mercury in contact with the solution depends upon this electromotive force, and becomes a maximum when this is zero. If, then, the difference of potential between the mercury and the solution be varied by any means and in either direction, the surface tension diminishes.

The mercury in a long tube (Fig. 185) will not flow through the capillary part at the lower end, unless the pressure of the mercury exceeds a certain amount, depending upon the surface tension of the mercury and the diameter of the tube, and since the drawn-out end of the tube tapers slightly, there will be an equilibrium position of the mercury meniscus for each value of the surface tension. The pressure of the mercury may be adjusted by varying the height of the column, by means of an arrangement not shown in the diagram, until the meniscus is in the field of the microscope, as shown at A. For an increase in the surface tension the meniscus will rise, while for a decrease it will sink,

[1] G. Lippmann, *Ann Chim. Phys.*, V. 5, p. 494. 1875.

and its position may therefore be adjusted by varying the resistance in the box R and therefore the difference of potential between the mercury in the capillary tube and that in the beaker. This difference of potential is adjusted until the meniscus reaches its highest point, or rather lowest, as seen inverted in the field of vision of the microscope, as at A, when the dilute sulphuric acid

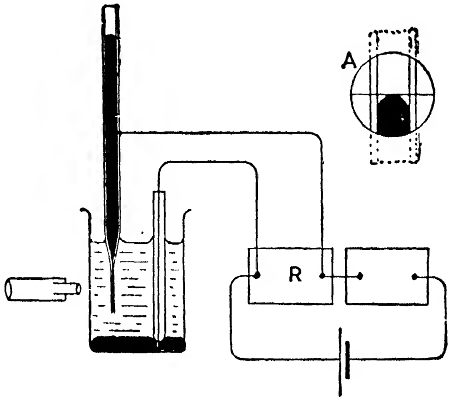


FIG. 185.

and the mercury in the capillary tube will be at the same potential. The potential difference as measured by the current in the resistance box R is equal to the electromotive force that exists between the mercury and the solution in the beaker.

A very convenient form of capillary electrometer is shown in Fig. 186. The capillary tube is of uniform bore and is nearly horizontal. Its lower portion contains mercury and its upper

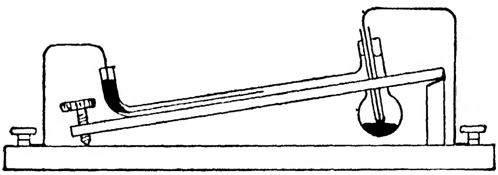


FIG. 186.

portion, together with the vessel, contains dilute sulphuric acid. The left-hand terminal is naturally positive to the solution, and if this difference of potential be diminished, the surface tension of the mercury surface in contact with the solution is increased, and the meniscus travels down the tube, reaching a limiting position when this difference of potential vanishes. If the difference of potential be further changed in the same direction, the meniscus travels back again. The sensitiveness of this electro-

meter may be increased by making the inclination of the tube to the horizontal less, by means of the screw, and by reading the position of the meniscus by means of a microscope.

Changes of p.d. of the order of 0·001 volt may be detected by means of this apparatus.

Secondary Cells or Accumulators.—In principle, any form of reversible cell may be used to store energy, for when a current is driven through the cell in the reverse direction, deposition will occur at the natural anode while the natural cathode is dissolved. The actions are reversed when the cell is used to produce current and the stored energy is recovered, at least in part.

The most generally satisfactory modern "accumulators" or secondary cells are still based on the type produced by Planté in 1859, using lead electrodes in a solution of sulphuric acid. After a preliminary "forming" process, in which current is sent repeatedly through the cell in opposite directions, the charging process leaves one plate (the negative) as lead, but the other is oxidised to the lead oxide PbO_2. During discharge, hydrogen ions, having delivered their positive charges to this latter plate, reduce it to a lower oxide and in the presence of the acid this presumably turns into lead sulphate ($PbSO_4$). Oxidation at the other plate also leads to formation of sulphate. When this coating of both plates with sulphate reaches a certain stage, the increasing similarity of the plates leads to a rapid fall in E.M.F. from the 2·1 volts characteristic of the fully charged cell, and the cell should be recharged.

Using e for the charge carried by a positive ion, the process of charging may be roughly represented as follows:—

Positive plate.

$$PbSO_4+SO_4^{--}+2H_2O=PbO_2+2H_2SO_4-2e.$$

Negative plate.

$$PbSO_4+2H^+=Pb+H_2SO_4+2e.$$

For discharge we have—

Positive plate.

$$PbO_2+H_2SO_4+2H^+=PbSO_4+2H_2O+2e.$$

Negative plate.

$$Pb+H_2SO_4+O^{--}=PbSO_4+H_2O-2e.$$

Hence during charge the electrolyte gains sulphuric acid. In fact, the density of the solution is the most convenient index of the condition of the cell; when fully charged it should not rise above the figure given by the makers, usually about 1·21, and the

discharge should stop before the density falls below 1·17, since the lead sulphate then becomes insoluble: the cell is "sulphated." The permissible density range is higher (1·1 to 1·25) for some cells.

In the later stages of charge, water is lost by electrolysis; hence the acid level is restored from time to time by the addition of distilled water (*not* acid). With a cell in good condition, the E.M.F. should remain very steady at about 2·1 volts for a very large fraction of the discharge. The internal resistance of a large cell is extremely low (of order 0·01 ohm or even less) and the type used for car starter batteries will supply currents of the order of hundreds of amperes for a few seconds at a time.

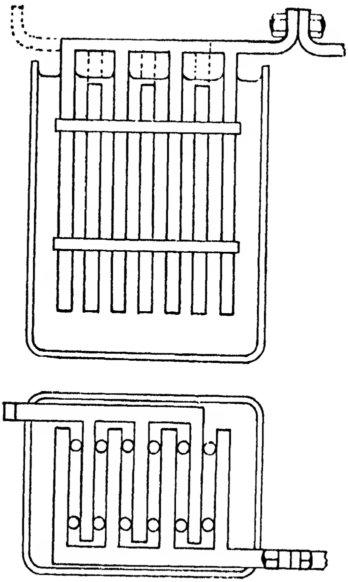


FIG. 187.

Modern cells often use "paste" plates (following a suggestion of Faure), with a mixture of lead oxides held in a lead grid: these are less durable but cheaper than plates "formed" from solid lead as in Planté's process. To obtain large exposed areas of active material and to keep the internal resistance low, the positive and negative plates are usually interleaved (Fig. 187) with insulating distance pieces between them. In the case illustrated, lengths of glass tubing are used for this purpose.

There are also on the market alkali cells in which the electrolyte is an aqueous solution of a mixture of alkali hydroxides and the plates are metal structures containing, in pockets, hydroxides of nickel and of iron or of cadmium. The E.M.F. of these cells is about 1·4 volts and is less steady than that of a lead cell, but the cell is less easily damaged by overloading or by being over-discharged or left standing. The electrolyte is slowly weakened by the action of atmospheric carbon dioxide and must be renewed from time to time.

CHAPTER VII

THERMO-ELECTRICITY

THE study of reversible thermo-electric effects dates from the discovery by Seebeck,[1] in the year 1826, that a current flows in a circuit consisting of two different metals when a difference of temperature is maintained between the two junctions. He arranged 35 metals in a series such that, when any two comprise a circuit, the current flows across the hot junction from the metal occurring earlier to that occurring later in the series. Seebeck's list comprises Bi—Ni—Co—Pd—Pt—U—Cu—Mn—Ti—Hg—Pb—Sn—Cr—Mo—Rh—Ir—Au—Ag—Zn—W—Cd—Fe—As—Sb—Te, and several others of doubtful composition, such as brass, commercial copper etc.

The discovery of the complementary phenomenon, the heating or cooling of a junction when a current flows across it, is due to Peltier,[2] who found in 1834 that on passing a current across a junction from bismuth to antimony, heat is absorbed at the junction, which is therefore cooled, but on reversing the direction of the current heat is developed, and the junction is warmed.

The Seebeck and Peltier phenomena may both be explained if we assume an electromotive force to exist at the junction of the two metals, its direction being from bismuth to antimony across the junction. If the circuit (Fig. 188) could be completed without the introduction of any further electromotive force, a current would flow in the direction BADC, a fall of potential occurring in the external circuit from A to B. The heat is produced by the current in the external circuit at the expense of the energy at the junction. Such a circuit, having only one junction, is impossible, and if the ends of the antimony and bismuth rods be bent round and brought into contact at a second point, the electromotive force at this junction is, of course, equal and opposite to that at the first, so that the resultant electromotive

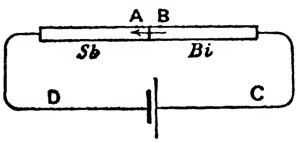


FIG. 188.

[1] T. J. Seebeck, *Pogg. Ann.*, Bd. VI., 1826.
[2] Peltier, *Ann. d. Chim. et de Phys.*, 2 Serie, **56**. 1834.

force in the circuit is zero, and there will be no current. If, however, a cell be introduced, so that a current is driven in the direction DCBA, it is found that the junction AB is cooled, and from our reasoning on p. 55 we should conclude that the direction of the Peltier electromotive force is from B to A. On reversing the cell, so that the current flows in the direction ABCD, the junction AB is warmed, which fact again indicates the presence of the electromotive force at the junction in the direction B—A.

On constructing a simple circuit of two metals—antimony and bismuth will do very well for our present purpose—we have seen that, owing to the opposition of the two equal electromotive forces at the junctions, there is no current. If, however, the junctions are not at the same temperature, these opposing electromotive forces are not necessarily equal; in fact, they are generally unequal, and the resultant electromotive force equal to their difference will maintain a current.

Let π_2 be the electromotive force at junction 2 (Fig. 189), at the higher temperature, and π_1 that at 1, at the lower temperature. In the case of the above metals $\pi_2 > \pi_1$, and if these are the only electromotive forces in the circuit, the resultant E.M.F. is $\pi_2 - \pi_1$, and the current is anti-clockwise.

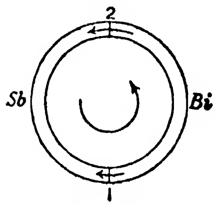


FIG. 189.

It will be seen that the current itself will cause a cooling at 2 and heating at 1, and we may therefore look upon the difference of temperature between the junctions as the condition for the current to flow, and further, the current flows until it has brought the circuit to uniform temperature.

It is not difficult to demonstrate the heating or cooling at an Sb–Bi junction [1] by placing a bismuth bar between two bars of antimony, and passing a current through the three in series as in Fig. 190. Two pieces of silk-covered, fine copper wire are wound one upon each half of the rod of bismuth; the two are connected to the two gaps of a metre bridge, and a balance found in the ordinary way. If, now, a current be passed through the rods from left to right, the junction A is warmed and B is cooled, and the two copper coils will now be at different temperatures. Since the electrical resistance of copper varies very rapidly with change of temperature, the balance of the metre bridge is now destroyed, and a galvanometer deflection will be observed, which deflection may be

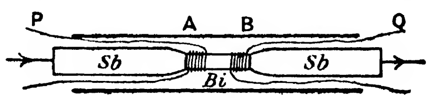


FIG. 190.

[1] S. G. Starling, *Nature*, Feb. 16, 1911.

reversed by sending the current through the rods from right to left. Peltier placed the junctions in glass bulbs, and observed the heating and cooling by the expansion or contraction of the air in the bulbs; but the amount of heat developed or lost is in all cases very small, so that atmospheric disturbances of temperature are liable to hide the effect sought. He also used a thermal junction, thus employing the Seebeck phenomenon, but it is desirable if possible to use an independent heat phenomenon for demonstrating the effect.

The Peltier effect must not be confused with the Joule production of heat. The latter, for a conductor of constant resistance r, is i^2r ergs per second, and is irreversible; that is, electrical energy is always converted into heat, the reverse process being impossible. Also the heating effect is proportional to the square of the current, and is therefore independent of its sign and direction, whereas the Peltier effect varies as the first power of the current and so depends upon its direction. Thus if the current flowing across a junction causes an absorption of πi ergs per second, π is the Peltier electromotive force, usually called the Peltier coefficient. A current in the opposite direction causes a heating at the same rate. Owing to the fact that the heat liberated at a junction diffuses by conduction through the mass of the metals, it is not convenient to measure the Peltier effects by means of the heating or cooling due to a known current; the actual method of determination will be described later.

Measuring heat in electrical units we have: heat absorbed at any junction $= \pm \pi i t$, where t is the time for which the current flows, the sign depending on the direction of the current.

Laws of Addition of Thermal Electromotive Forces.—In measuring the electromotive force in any circuit due to thermoelectric effects, it is nearly always necessary to insert some piece of apparatus, such as a galvanometer, somewhere in the circuit, and since this generally involves the presence of more than the two original metallic junctions, it is important to formulate the laws according to which the electromotive forces produced by additional junctions may be added. There are two such laws.

1. *Law of Intermediate Metals.*—The insertion of an additional metal into any circuit does not alter the whole electromotive force in the circuit, provided that the additional metal is entirely at the temperature of the point of the circuit at which it is inserted.

This law may be taken as the result of experiment, but we may see that it follows from the second law of thermodynamics; for if a number of metals A, B, C etc. (Fig. 191), are joined in series to form a complete circuit, there is no current in the circuit when the temperature is everywhere the same. Should a

current flow, it will immediately cause heating or cooling at the junctions, and the energy required to maintain the current would be obtained by heating some parts of the circuit and cooling others, the divergence of temperature becoming greater the longer the current flows. As there is no chemical action in the circuit, the above process would be in contradiction to the second law of thermodynamics, and we therefore conclude that there is no current, and that the algebraic sum of the electromotive forces in the circuit is zero. The same reasoning would apply if C were removed, and therefore we conclude that the introduction of C when it is entirely at the temperature of the point at which it is inserted does not alter the total electromotive force in the circuit.

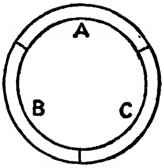


FIG. 191.

This is equally true though the junction between A and B at which C is inserted should be at some other temperature, for this does not affect the electromotive force occurring at the unaltered junction of A and B, this being determined by its own temperature only.

2. *Law of Intermediate Temperatures.*—The electromotive force for a couple with junctions at T_1 and T_3 is the sum of the electromotive forces of two couples of the same metals, one with junctions at T_1 and T_2 and the other with junctions at T_2 and T_3.

For, in Fig. 192 (α) let the electromotive force for the T_2–T_3 couple be $[e]_2^3$ and that for the T_1–T_2 couple be $[e]_1^2$. Then if the junctions at the temperature T_2 be placed in contact there is no change, because like metals at the same temperature only are joined. If then the junctions be opened to form the arrangement (β) there is again no change in the resultant electromotive force, for the two contacts destroyed both had the same Peltier effect π_B^A at temperature T_2, and these are directed oppositely in the compound circuit. We therefore conclude that

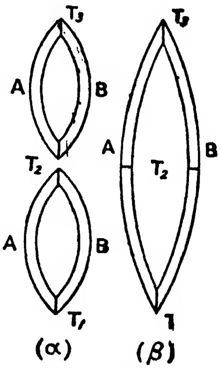


FIG. 192.

$$[e]_1^3=[e]_1^2+[e]_2^3.$$

Application of Thermodynamics.—Since the Peltier effect is a reversible one, a thermal couple is an arrangement for deriving useful energy by the absorption of heat at one temperature, part of which is given back at a lower temperature, the difference in the amount absorbed and that given up being the energy applicable for external purposes. Thus the current may be used for

driving an electro-motor in which case the energy takes the form of mechanical work. Although the available energy is usually very small in amount this does not vitiate our argument.

Lord Kelvin [1] pointed out that, the processes being entirely reversible, the arrangement is in reality a heat engine with source at one temperature (T_2) and refrigerator at a lower temperature (T_1), and that the ratio of the heat absorbed at temperature T_2 to that given up at T_1 should be the same as that of T_2 to T_1, where T_2 and T_1 are absolute temperatures.

Now, on carrying a charge q round the circuit, the heat absorbed at the hot junction is $\pi_2 q$, measured in absolute units, and that given up at T_1 is $\pi_1 q$,

hence, $$\frac{\pi_2 q}{\pi_1 q}=\frac{T_2}{T_1}, \text{ or, } \frac{\pi_2}{\pi_1}=\frac{T_2}{T_1},$$

and therefore, $$\frac{\pi_2-\pi_1}{\pi_1}=\frac{T_2-T_1}{T_1}.$$

Now, if $\pi_2-\pi_1=e$, the whole electromotive force in the circuit—

then, $$e=\pi_1\left(\frac{T_2-T_1}{T_1}\right).$$

It would therefore follow that if one junction is maintained at constant temperature T_1, then π_1 is constant, and $e \propto (T_2-T_1)$. Now it may easily be shown that this is not true; for if a piece of copper and a piece of iron wire be twisted together at one end and the other ends connected to a galvanometer, it will be found on heating the copper-iron junction with a burner, that the resulting current, and therefore electromotive force, increases at first, then diminishes and, passing through zero, actually becomes reversed.

Obviously then, e is not proportional to T_2-T_1.

Lord Kelvin (then Prof. Wm. Thomson) therefore concluded that the Peltier effect was not the only source of electromotive force in the circuit, and pointed out the likelihood of another, existing between the different parts of a metal at different temperatures.

If then, for any substance, σ is the electromotive force due to unit difference of temperature between two points of it, $\int_{T_1}^{T_2}\sigma dT$ is the total electromotive force between points at temperatures T_1 and T_2, and taking σ_a and σ_b for the value of σ in the two metals A and B in Fig. 193, our equation of electromotive forces for the whole circuit will now become

[1] W. Thomson, *Phil. Trans. Roy. Soc.*, 1855.

$$e=\pi_2-\pi_1-\int_1^2\sigma_a dt+\int_1^2\sigma_b dt,$$

the small arrows indicating the directions in which the various electromotive forces tend to drive the current round the circuit. σ is assumed to be positive for both metals, that is, the electromotive force is directed from points of lower to those of higher temperature, and could the circuit be completed by a neutral conductor so that this is the only electromotive force in the circuit, the current would flow in the external circuit from points of high to points of low temperature.

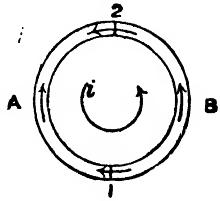


FIG. 193.

The quantity σ is called the Thomson coefficient, and the existence of the electromotive force involves an absorption of heat if a current flows in the direction of the electromotive force, since its direction is such that the electromotive force tends to maintain the current, and therefore to give energy to the circuit, which energy is supplied at the expense of the heat of the metal itself. If the current be reversed, heat is liberated for a corresponding reason (p. 55). The sign of σ may be positive or negative, which means that the Thomson electromotive force may act in such a direction that it tends to drive the current in the external part of the circuit from points of high to points of low temperature or *vice versâ.* Thus, if σ is positive the state of affairs is shown in Fig. 194 (i), where the ordinates indicate the temperature, and the small arrows the Thomson electromotive force. The current passing in the direction ABC absorbs heat in AB, since it is flowing in the direction of the electromotive force; that is, from points of lower to those of higher temperature; and for a corresponding reason it gives out heat in BC, just as a flow of an ordinary liquid down a tube heated at B would do. On the other hand, if σ is negative, we have the condition shown in Fig. 194 (ii). A current flowing from A to C would give out heat in the part AB and absorb heat in BC; that is, it gives out heat when flowing from colder to hotter points, and absorbs it when flowing from hotter to colder points, so that if we wish to find any analogy with the flow of liquid in a tube heated in the middle, we must imagine the liquid to have a negative specific heat.

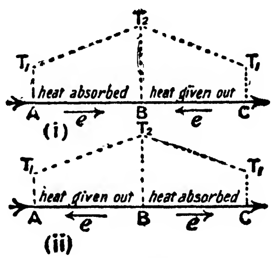


FIG. 194.

σ is positive for the metals Cd, Zn, Ag, Cu, and negative for Fe, Pt and Pd.

The Thomson effect may be exhibited [1] in a manner similar to that for the Peltier effect, but in this case the difficulty is greater, on account of the fact that a considerable temperature gradient is necessary for the exhibition of the effect, and thus it is not easy to measure the small additional reversible heating and cooling due to the current. If, however, an iron rod be bent into the shape shown in Fig. 195 with the two limbs, on which the resistance coils P and Q are wound very close together and packed round with asbestos wool, then if P and Q are placed in gaps of a metre bridge as before and a balance found, a current of 10 amperes flowing round ABC will cause a disturbance in the balance when B is heated to red heat with a bunsen burner, and A and C immersed in mercury baths. In this way a very steep temperature gradient in BA and BC can be maintained, and the change of resistance in P and Q, due to the current, occurs in a manner which shows that heat is given out when the current flows up the temperature gradient, and the limb in which the current flows down the temperature gradient is cooled.

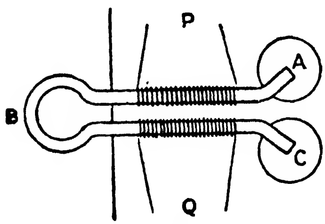


FIG. 195.

The effect with copper is in the reverse direction, and is much smaller, both on account of the smallness of the Thomson coefficient σ, and the difficulty of maintaining sufficient temperature gradient owing to the high thermal conductivity of the metal.

Thermo-electric Power.—The equation of electromotive force—

$$e=\pi_2-\pi_1-\int_1^2\sigma_a dT+\int_1^2\sigma_b dT,$$

may be written in the form—

$$e=\int_1^2\frac{d\pi}{dT}dT-\int_1^2(\sigma_a-\sigma_b)dT,$$

$\frac{d\pi}{dT}$ being the rate of change with temperature of the Peltier coefficient for the two metals, and π_2 and π_1 the upper and lower limits of the integral $\int\frac{d\pi}{dT}dT$.

Or, again,

$$e=\int_1^2\left\{\frac{d\pi}{dT}-(\sigma_a-\sigma_b)\right\}dT.$$

[1] S. G. Starling, *Nature*, Feb. 16, 1911.

Differentiating this equation with respect to T, we have—

$$\frac{de}{d\mathrm{T}}=\frac{d\pi}{d\mathrm{T}}-(\sigma_a-\sigma_b).$$

$\frac{de}{d\mathrm{T}}$ is called the *Thermo-electric power* for the two metals, and is the rate of change of the electromotive force acting round a couple with change of temperature of one junction. The thermo-electric effects in a circuit may be very conveniently represented on a diagram in a manner suggested by Prof. Tait,[1] the values of $\frac{de}{d\mathrm{T}}$, the thermo-electric power, being plotted against the temperature. Then, at a temperature represented by the point F in Fig. 196, the thermo-electric power P is represented by EF, and the thickness of the strip being dT,

$$\text{area of strip}=\frac{de}{d\mathrm{T}}\,.\,d\mathrm{T}=\mathrm{P}d\mathrm{T}=de.$$

Hence the area of the strip represents the electromotive force acting round the couple, the difference of temperature of the junctions being dT.

By the law of intermediate temperatures, the electromotive force round a couple having junctions at temperatures repre-

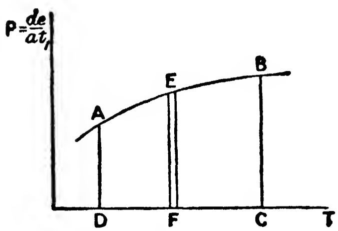


FIG. 196.

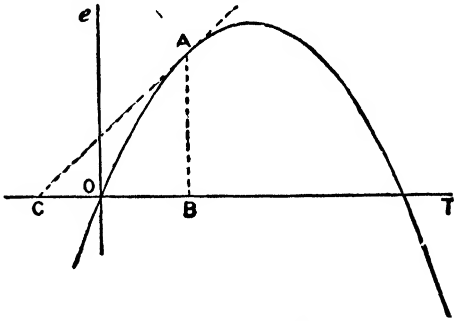


FIG 197.

sented by D and C respectively is equal to the sum of the electromotive forces for a number of couples having differences of temperature dT, provided that the first junction has temperature corresponding to D and the last to C. Thus, the electromotive force is equal to the sum of the areas of strips, such as EF, and this is the area of the whole figure ADCB.

The electromotive force for finite differences of temperature may be found by experiment, by measuring the total electro-

[1] Prof. Tait, *Proc. Roy. Soc. Edin.*, p. 597. 1871.

motive force round a circuit, when one junction is maintained at constant temperature and that of the other varied. The shape of the curve usually obtained is shown in Fig. 197, and the rate of increase of e, when the temperature of the hot junction is represented by OB is the ratio $\frac{\text{AB}}{\text{BC}}$, where AC is a tangent to the curve at the point A.

Thus, $$\frac{de}{d\text{T}}=\frac{\text{AB}}{\text{BC}}=\text{P},$$

and the curve (Fig. 196) may now be plotted for P, the thermo-electric power as derived from Fig. 197.

The experimental methods of measuring the electromotive force due to a couple will be considered later, but we may note that the E.M.F.–Temperature curves for most metals approximate to parabolas, which would lead to the (thermo-electric power)-(temperature) curves being straight lines, as we shall see on p. 215.

Second Law of Thermodynamics.—In applying the second law of thermodynamics to the couple, we must now take the Thomson effect into account, since the heat is not all absorbed at the hot junction nor all given up at the cold junction. Suppose a charge q to pass round the couple consisting of metals A and B, having its junctions at temperatures T_2 and T_1 (Fig. 198). As before (p. 206), when charge q passes round the circuit in an anti-clockwise direction in the figure, the heat absorbed at the hot junction is $q\pi_2$, and that given out at the cold junction, $q\pi_1$, that given out in passing through the metal A is $q\int_1^2 \sigma_a dt$, and that absorbed in passing through B is $q\int_1^2 \sigma_b dt$ (see Fig. 193).

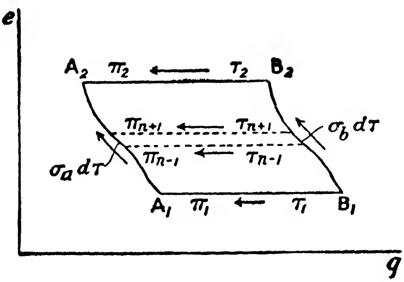


FIG. 198.

Let us consider the couple to consist of an infinitely great number of small couples at temperatures, T_{n-1} and T_{n+1}, etc., varying from T_1 to T_2. From the law of intermediate temperatures (p. 205), the electromotive force for the whole couple is the sum of the electromotive forces for the small couples. For the small couple whose junctions are at T_{n+1} and T_{n-1}, bearing in mind the direction of the current and the fact that when this is

in the direction of the Peltier or Thomson electromotive force, heat is absorbed and *vice versâ*, we see from Fig. 198 that—

heat absorbed at $T_{n+1}=q\pi_{n+1}$,
heat given out at $T_{n-1}=q\pi_{n-1}$,
heat given out in metal A at mean temp. $T_n=q\sigma_a dT$,
heat absorbed in metal B at mean temp. $T_n=q\sigma_b dT$.

Since all these processes are reversible, we can apply the second law of thermodynamics to the cycle, which states that $\Sigma\frac{Q}{T}=0$.

Thus, for the elementary cycle—

$$\frac{q\pi_{n+1}}{T_{n+1}}-\frac{q\pi_{n-1}}{T_{n-1}}-\frac{q\sigma_a dT}{T_n}+\frac{q\sigma_b dT}{T_n}=0.$$

For the adjacent small cycle we have—

$$\frac{q\pi_{n-1}}{T_{n-1}}-\frac{q\pi_{n-3}}{T_{n-3}}-\frac{q\sigma_a dT}{T_{n-2}}+\frac{q\sigma_b dT}{T_{n-2}}=0,$$

where the lower temperature for one cycle is the upper temperature for the next, and so on. Adding up these equations for all the elementary cycles, remembering that the first has upper temperature T_2, and the last, lower temperature T_1, the terms $\frac{q\pi_{n+1}}{T_{n+1}}$, etc., all cancel out except the first and the last.

$$\therefore\ \frac{q\pi_2}{T_2}-\frac{q\pi_1}{T_1}-q\int_1^2\frac{\sigma_a dT}{T}+q\int_1^2\frac{\sigma_b dT}{T}=0,$$

$$\text{or, } \frac{\pi_2}{T_2}-\frac{\pi_1}{T_1}-\int_1^2\frac{\sigma_a-\sigma_b}{T}dT=0.$$

Now $\frac{\pi_2}{T_2}$ and $\frac{\pi_1}{T_1}$ are the limits at T_2 and T_1 of the integral $\int\frac{d}{dT}\left(\frac{\pi}{T}\right)dT$,

$$\therefore\ \int_1^2\frac{d}{dT}\left(\frac{\pi}{T}\right)dT-\int_1^2\frac{\sigma_a-\sigma_b}{T}dT=0.$$

Differentiating this, we have—

$$\frac{d}{dT}\left(\frac{\pi}{T}\right)-\frac{\sigma_a-\sigma_b}{T}=0.$$

$$\therefore\ \sigma_a-\sigma_b=T\frac{d}{dT}\left(\frac{\pi}{T}\right).$$

Substituting this value of $\sigma_a-\sigma_b$, which is a consequence of the

application of the second law of thermodynamics, in equation—

$$\frac{de}{dT}=\frac{d\pi}{dT}-(\sigma_a-\sigma_b), \qquad \text{(p. 209)}$$

we have,

$$\frac{de}{dT}=\frac{d\pi}{dT}-T\frac{d}{dT}\left(\frac{\pi}{T}\right)$$

$$=\frac{d\pi}{dT}-T\left\{\frac{1}{T}\cdot\frac{d\pi}{dT}-\frac{\pi}{T^2}\right\}$$

$$=\frac{\pi}{T},$$

From which,

$$\pi=T\frac{de}{dT}.$$

Thus the Peltier coefficient for the junction of a pair of metals is the product of the absolute temperature of the junction (T) and the rate of change of the electromotive force for the whole circuit with change of temperature of the junction $\left(\frac{de}{dT}\right)$.

Thermo-Electric Diagram.—Let us now return to the consideration of the thermo-electric diagram; we shall find that the relation $\pi=T\frac{de}{dT}$ enables us to interpret the diagram more fully than we could hitherto. If A_1A_2 (Fig. 199) is the line whose ordinates are the thermo-electric powers P at different temperatures, of the metal A, with respect to the metal B, the area $A_2A_1B_1B_2$ represents the electromotive force acting round the couple when the temperatures of the junctions are T_1 and T_2. These, on the diagram, are imagined to be measured from the absolute zero of temperature, the direction of the effective electromotive force round the circuit being in the order of the letters, that is anti-clockwise in the diagram.

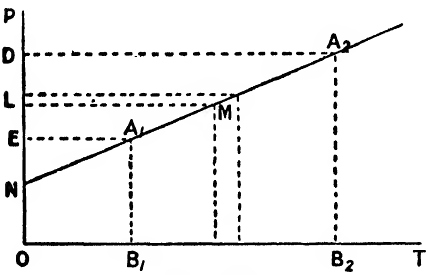


FIG. 199.

$$\text{Now } \pi_2=T_2\frac{de}{dT}=T_2P_2.$$

But $A_2B_2=P_2$, and $OB_2=T_2$,

$\therefore \pi_2$=area of rectangle B_2A_2DO.

Similarly, π_1=area of rectangle B_1A_1EO.

Again, from equation $\sigma_a-\sigma_b=T\frac{d}{dT}\left(\frac{\pi}{T}\right)$ on p. 211;

since,
$$\frac{\pi}{T}=\frac{de}{dT}=P,$$

$$\sigma_a-\sigma_b=T\frac{dP}{dT},$$

and,
$$(\sigma_a-\sigma_b)dT=TdP.$$

But the area of the strip LM is TdP, because $LM=T$, and width of strip $=dP$.

$$\therefore \int_1^2(\sigma_a-\sigma_b)dT=\text{area } A_1A_2DE.$$

We can therefore identify all the thermal electromotive forces acting round the couple, as areas upon the diagram.

Thus,
$$\pi_2=\text{area } B_2A_2DO,$$
$$\pi_1=\text{area } B_1A_1EO,$$

and,
$$\int_1^2(\sigma_a-\sigma_b)dT=\text{area } A_1A_2DE.$$

Which would, from our electromotive force equation—

$$e=\pi_2-\pi_1-\int_1^2(\sigma_a-\sigma_b)dT,$$

give,
$$e=\text{area } B_2A_2A_1B_1.$$

It will be seen that the Peltier and Thomson effects are electromotive forces which would drive a current from points whose positions upon the thermo-electric diagram are lower to those which are higher.

The electromotive force round a couple may be represented as a function of the thermo-electric power, or the Peltier coefficient for—

$$P_A=\frac{de}{dT}=\frac{\pi}{T}.$$

$$\therefore\ de=P_AdT=\left(\frac{\pi}{T}\right)dT,$$

$$\text{and, } [e]_1^2=\int_1^2P_AdT=\int_1^2\left(\frac{\pi}{T}\right)dT,$$

where $[e]_1^2$ is the electromotive force acting round a couple A–B the temperatures of whose junctions are T_1 and T_2 respectively, and P_A is the thermo-electric power of A with respect to B.

In the above reasoning one metal (B) has been taken as a standard and the thermo-electric powers of the other metal (A)

plotted with respect to it. If now a third metal (C) be combined with B to form a couple, we should have an exactly similar set of relations for C and B. Thus for temperatures of junctions T_1 and T_2 (Fig. 200)—

$$e = \text{area } C_2C_1B_1B_2$$
$$\pi_2 = T_2 \,.\, B_2C_2, \text{ and, } \pi_1 = T_1 \,.\, B_1C_1, \text{ etc.,}$$

and it follows from the law of intermediate metals that for the couple made up of A and C—

$$e = \text{area } A_1A_2C_2C_1$$
$$\pi_2 = T_2 \,.\, A_2C_2 \text{ and } \pi_1 = T_1 \,.\, A_1C_1.$$

Hence if any one metal be taken as standard and the thermo-electric powers of a number of others be plotted with respect to it, the thermo-electric powers with respect to each other of the different metals will simply be the difference of the respective ordinates for the thermo-electric lines of the two metals. And further, if it is desired to change the standard metal, all that is necessary is to replot the curves, with the differences of the ordinates measured from the new standard as ordinates upon the new diagram. This will not change the relations of the thermo-electric powers of the different metals, nor will it alter any of the areas in the diagram, and the respective electromotive forces also will be unchanged. Thus, in Fig. 200, if we subtract A_1B_1, A_2B_2, etc., from all the appropriate ordinates, A_1A_2 will now be horizontal, B_1B_2 will slope downwards from left to right, and the slope of C_1C_2 will be increased, but the electrical quantities involved will all be unchanged.

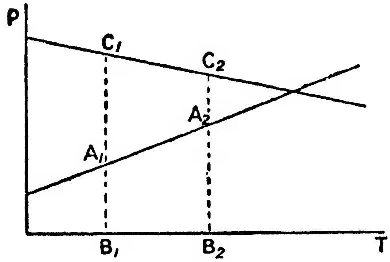


FIG. 200.

The metal usually taken as a standard with respect to which the thermo-electric powers of the others are plotted is lead, the reason being that the Thomson coefficient for lead is supposed to be zero,[1] but should this subsequently prove not to be the case, the usefulness of the diagram would not be affected, and we could, if we chose, knowing the value of σ for lead, replot the diagram, taking an ideal metal for which $\sigma=0$ as standard.

If we assume that for lead $\sigma=0$, which is certainly very nearly true, and in Fig. 199 take B to be lead, equation $\sigma_a - \sigma_b = T\frac{dP}{dT}$ becomes $\sigma_a = T\frac{dP}{dT}$, since $\sigma_b = 0$.

[1] Le Roux, *Ann. de Chim. et de Phys.*, 4 serie, 10, p. 201. 1867.

Thus since T is essentially positive, the Thomson coefficient σ has the same sign as $\frac{dP}{dT}$, and if P increases with the temperature σ is positive, if P decreases with rising temperature σ is negative. It will then be seen from the diagram (Fig. 203) that for cadmium, zinc etc., σ is positive, and for iron, palladium etc., it is negative, and further, that since $\frac{dP}{dT}=\tan\theta$, where θ is the inclination of a thermo-electric line to the axis at any point—

$$\sigma=T\tan\theta.$$

Neutral Temperature.—So far we have made no assumption as to the shape of the thermo-electric lines; but if they are straight lines we can then calculate the electromotive force round any couple when we know the equations to the thermo-electric lines.

In Fig. 201, let $P_a=m_aT+c_a$, and $P_b=m_bT+c_b$ be the equations

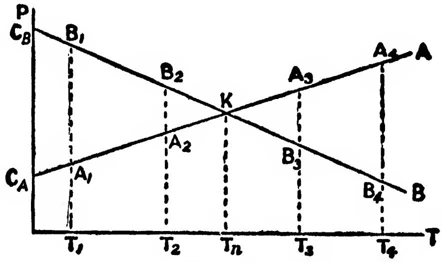


FIG. 201.

to the thermo-electric lines for the metals A and B. Then—

$$[e]_1^2=\int_1^2(P_a-P_b)dT=\int_1^2[(m_a-m_b)T+c_a-c_b]dT,$$

which, when integrated gives—

$$e=\tfrac{1}{2}(m_a-m_b)(T_2^2-T_1^2)+(c_a-c_b)(T_2-T_1).$$

If the temperature T_1 be fixed while T_2 be varied, we see that the equation connecting e and T_2 is that of a parabola. It may be written in the form—

$$e=(T_2-T_1)\left\{\left(\frac{T_2+T_1}{2}\right)(m_a-m_b)+(c_a-c_b)\right\}.$$

Hence e is zero when $T_2=T_1$, which would be expected; but it again becomes zero when

$$\frac{T_2+T_1}{2}=-\frac{c_a-c_b}{m_a-m_b}.$$

That is, when the average temperature of the junctions is

$-\frac{c_a-c_b}{m_a-m_b}$. This temperature is called the *Neutral Temperature*, and is that at which $P_a=P_b$, or the thermo-electric lines, $P_a=m_aT+c_a$, and $P_b=m_bT+c_b$, intersect. Calling it T_n, we may now write the electromotive force equation—

$$e=(m_a-m_b)(T_2-T_1)\left\{\frac{T_1+T_2}{2}-T_n\right\}.$$

T_n is evidently the temperature corresponding to the point E (Fig. 202) on the E.M.F.–Temperature diagram at which the curve ceases to rise and begins to descend, for on referring to Fig. 201 it will be seen that with one junction at fixed temperature T_1, the area $B_1A_1A_2B_2$ which represents the electromotive force round the couple, increases as T_2 rises, until the neutral point K is reached. When this is passed, as at temperature T_3, the area A_3KB_3 must be deducted from the area B_1A_1K to get the effective electromotive force round the couple; for at temperatures below T_n the thermo-electric power of B is greater than that of A, but this is reversed when T_n is passed, the thermo-electric power of B with respect to A at T_n being zero. Below T_n the Peltier and Thomson effects have been so directed in A_1A_2, A_2B_2, B_2B_1, that they tend to drive a current round the circuit from A to B across the hot junction, but at T_3 the effect in B_3A_3 is such that it tends to drive the current from B to A at the hot junction, and its value grows until a temperature T_4 is reached, which is as much above T_n as T_1 is below it, when area $A_1KB_1=B_4A_4K$, and the resultant electromotive force round the couple is zero. This temperature corresponds to the point F in Fig. 202. The student may verify the fact that in Fig. 201—

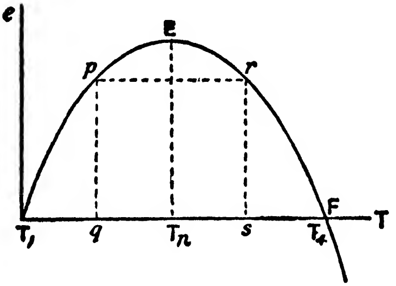


Fig. 202.

$$\text{area }(A_4B_4\times T_4)-\text{area}\int_1^4\sigma_a dT+\text{area}(A_1B_1\times T_1)-\text{area}\int_1^4\sigma_b dT=0.$$

If now the temperature of the hot junction be raised above T_4 the resultant electromotive force in the couple is reversed. For this reason T_n is sometimes called the temperature of inversion, as the direction of the resultant electromotive force changes sign as the average temperature of the couple passes this value.

Results of Experiments.—A curve such as that shown in

Fig. 202 is easy to determine experimentally, and from it the neutral temperature can be accurately found. The temperature corresponding to the highest point E can only be read approximately, since at this point the electromotive force is changing very slowly with temperature. But by taking two points p and r for which the electromotive force is the same, the mid-point between q and s is the neutral temperature T_n. If this be done for some metal A, using lead for the metal B, $\sigma_b=0$, $m_b=0$, and $c_b=0$, since the thermo-electric line for B is now the temperature axis, and therefore $T_n=-\frac{c_a}{m_a}$.

Also when one junction is at the neutral temperature (say $T_1=T_n$), remembering that $m_b=0$, we have, $e=\frac{1}{2}m_a(T_2-T_n)^2$, either from the diagram (Fig. 199) taking T_n as the intersection of A_1A_2 with the axis of T, or by substitution in the equation for e on p. 216. Thus if the electromotive force for any chosen temperature be noted from the curve (Fig. 202), m_a can be calculated, and since $T_n=-\frac{c_a}{m_a}$, c_a can also be found, and the equation—

$$P_a=m_aT+c_a$$

for the thermo-electric line becomes known.

Referring to equation $\sigma=T \tan \theta$ (p. 215), we see that when the thermo-electric lines are straight, $\tan \theta$ is identical with m_a, which is constant, and that σ is therefore proportional to the absolute temperature.

If it is inconvenient to compare the given metal directly with lead, the electromotive forces for a couple made up of the metal with some other which has previously been compared with lead may be found, and the electromotive forces for the metal and lead found by means of the law of intermediate metals.

Tait [1] found that for most of the metals the E.M.F.–Temperature curves are approximately parabolas, and therefore the thermo-electric lines are straight; but exceptions occur, as in the cases of iron and nickel, which exhibit several points of inflection at high temperatures. The thermo-electric line for iron cuts that for an alloy of platinum and iridium at several high temperatures, and it is pointed out that if an Fe–(Pt-Ir) couple have one junction at each of two temperatures at which the thermo-electric lines intersect, a current will be maintained on account of the Thomson effect alone, for the Peltier coefficients at these points are zero. The curves of Fig. 203 are taken from Prof. Tait's paper, but the thermal electromotive forces are converted from his arbitrary units (the Grove cell, E.M.F. $=1{\cdot}7$ volt) approximately into microvolts.

[1] Prof. Tait, *Trans. Roy. Soc. Edin.*, p. 125. 1873.

The thermo-electric lines for iron at temperatures up to 1000° C. have been determined by G. Belloc,[1] and these show clearly

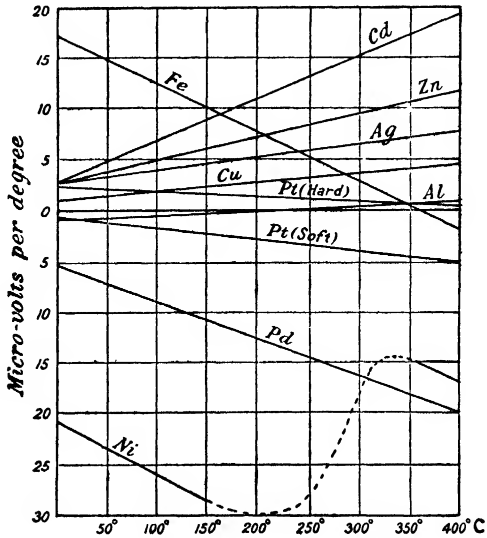


FIG. 203.

(Fig. 204) the points of inflection mentioned by Tait. The thermo-electric powers are given in microvolts per degree, with respect to the metal platinum, which was taken for reference. The effect of various percentages of carbon in the iron upon the thermo-electric power is also indicated by the dotted lines in the

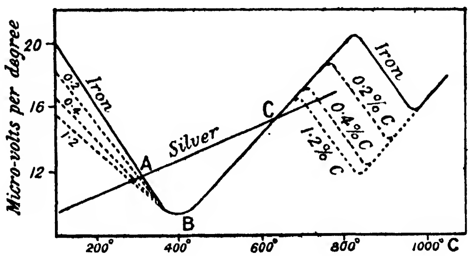


FIG. 204.

diagram. The approximate position of the thermo-electric line of silver with respect to platinum is also placed upon the same diagram, and it will be seen that if a silver-iron couple be constructed and the junctions maintained at 310° C. and 620° C. respectively, the Peltier coefficients at these temperatures are

[1] G. Belloc, *Ann. de Chim. et de Phys.*, **30**, p. 42. 1903.

zero, and a current will then be maintained on account of the Thomson effect alone, the effective electromotive force acting round the couple being represented by the area to scale of the figure ABC.

The following values of the E.M.F. in microvolts for a couple consisting of platinum with one of the following metals, one junction being at 0° C. and the other at either −190° C. or +100° C., are given by Kaye and Laby.

Metal	−190°	+100°	Metal	−190°	+100°
Aluminium .	+390	+380	Lead . . .	+210	+410
Bismuth . .	+12300	−6500	Nickel . .	+2220	−1640
Copper . .	−200	+740	Silver . . .	−140	+710
Gold . . .	−120	+730	Zinc . . .	−120	+750
Iron . . .	−2900	+1600	Constantan *.	—	−3440

* Eureka, 60 Cu, 40 Ni.

From these results, the constants in the E.M.F. temperature equation on p. 216 can be found, or the thermo-electric lines may be determined.

Experimental Measurements.—The electromotive force in a thermal couple may be measured by placing a calibrated galvanometer in the circuit and observing the current. Then, knowing the resistance of the circuit, the corresponding electromotive force may be found; but a much better way is to employ the potentiometer, as in this case the current in the couple is zero when a balance is obtained. Since the electromotive force is usually of the order of a few millivolts, the potentiometer must be modified so that it can measure much smaller E.M.F.'s than usual. A wire AB about a metre long (a metre bridge will do very well) is connected in series with a

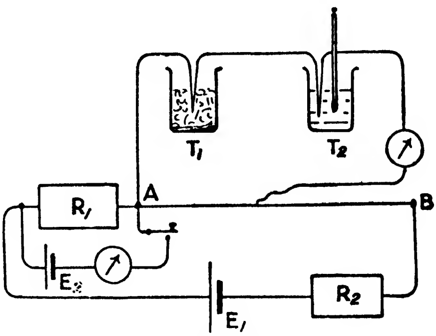


FIG. 205.

resistance box R_1 (Fig. 205), a rheostat R_2, and a secondary cell E_1. The resistance per centimetre of AB being known, the fall of potential in microvolts per centimetre of it can be found when the current is adjusted by means of R_2, so that the potential difference between the ends of R_1 is equal to the electromotive force of the standard cell E_2. Then, the junction T_1 being kept

in ice and water, the temperature of T_2 may be varied, and the electromotive force of the couple found from the length of wire AB necessary to produce a balance.

A more convenient apparatus for the same purpose, made by the Cambridge Instrument Company, is illustrated in Fig. 206. The secondary cell B maintains a steady current in the circuit $BVVSSR_2R_1B$. Using a standard cadmium cell C between M and N, the current in the main circuit is adjusted by means of the rheostats R_1 and R_2 until a balance is obtained, and SC and SN are arranged to be of such resistance that for the proper value of the main current, the fall of potential over 50 ohms of circuit is 1 volt. The coil MVC has 29 sections, each of resistance 0·05 ohm, and the fall of potential per section is thus 1 millivolt. VV is the slide wire upon which the final balancing is performed, and has a resistance of 0·06 ohm, so that the difference of poten-

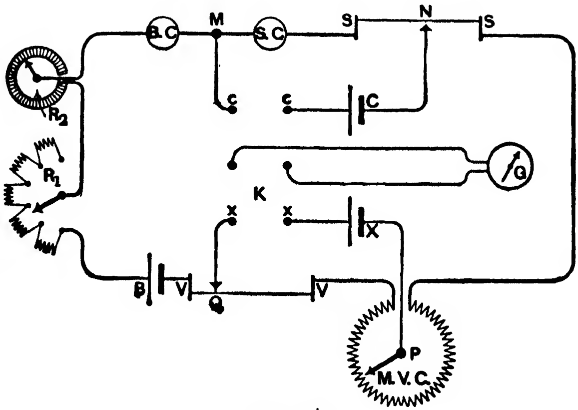


FIG. 206.

tial between P and Q due to the current may be varied at will from 0 to 30·2 millivolts. This is, therefore, the range of variation of the electromotive force of the thermal couple to be measured, the couple being placed at X. As the electromotive force of a thermal couple rarely exceeds 30 millivolts, the instrument is a very convenient one for rapidly calibrating such couples.

If a number of points on the E.M.F.–Temperature curve be obtained with one junction at fixed temperature, and the other variable, the apparatus forms a convenient pyrometer for measurements of temperature over considerable range.

Applications to Thermometry.—The electromotive force in a thermal couple, although very small, has, as a rule, a circuit of very low resistance in which to produce a current, which may therefore be considerable. One of the best-known applications is the detection of small amounts of radiant heat by means of

the *Thermopile.* The effect produced by one junction is multiplied by arranging a number in series. Antimony and bismuth bars alternate, one set of junctions A (Fig. 207) being exposed to the radiation, and the other set B being protected by a metal cap to maintain them at constant temperature.

A more sensitive arrangement is seen in Boys' radio-micrometer,[1] in which the couple and the galvanometer are combined in one instrument, the loop of wire which hangs between the poles of a powerful permanent horseshoe magnet terminating in a piece of antimony and one of bismuth soldered together at the tips. The radiation falling upon this junction warms it, and the thermo-electric force is established in the circuit, producing a current in the loop which, hanging in a magnetic field, experiences a couple. This arrangement has been modified to form a galvanometer by Duddell (see p. 75).

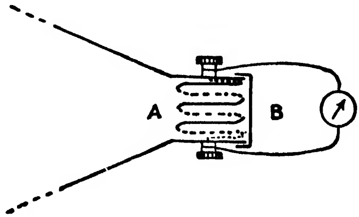


FIG. 207.

On referring to the thermo-electric diagram (Fig. 203) it will be seen that some of the thermo-electric lines, for example those of copper and silver, are nearly parallel; if they were actually parallel, the electromotive force round one of these couples would be proportional to the difference of the temperature of the junctions, since the figure $A_2C_2C_1A_1$ (Fig. 200) in this case becomes a parallelogram, and its area is proportional to the perpendicular distance between the sides A_1C_1 and A_2C_2, that is, to T_2-T_1. This is approximately true for some couples. For measurement of high temperatures, the couple is usually of pure platinum and an alloy of platinum and iridium, or platinum and rhodium, and is enclosed in a tube of suitable material for withstanding the temperature to which it will be exposed. In series with the couple, a millivoltmeter may be employed, which may be graduated in degrees Centigrade, and is of the type described on p. 81.

Thermo-milliammeter.—A sensitive form of ammeter, applicable to the measurement of small alternating or continuous currents, has been devised by Sir J. A. Fleming,[2] in which the heating produced by the current flowing in a fine constantan wire AB (Fig. 208) warms the junction of a tellurium-bismuth couple. The fine wires of tellurium and bismuth are soldered to the constantan wire, and the whole is situated in a high vacuum in a

[1] C. V. Boys, *Phil. Trans.*, 180, A., p. 159. 1889.
[2] J. A. Fleming, "The Principles of Electric Wave Telegraphy and Telephony."

glass vessel. In this way considerable sensitiveness is obtained, and the galvanometer G in series with the thermo-electric couple may be calibrated by passing a known continuous current through AB. Since the heating effect is proportional to the square of the current the instrument may be used when the current is alternating (p. 83).

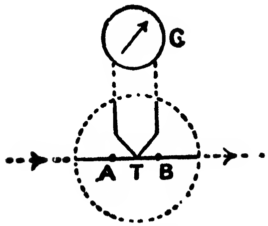


FIG. 208.

Radio-Balance.—The absorption of heat at a thermal junction, when the direction of the current is the same as that of the Peltier electromotive force, has been employed by Prof. Callendar[1] for the measurement of radiant heat. The radiation is absorbed by a blackened copper disc upon which it falls, and the rise in temperature in a given time might be measured by means of a thermo-electric couple of iron and constantan, which also acts as a suspension for the disc. To determine the rate of absorption of energy from the rate of rise of temperature would require a knowledge of the heat capacity of the system and the losses due to conduction and radiation, but, instead, the temperature is maintained constant by passing a current through a second thermal junction attached to the disc, and varying the strength until the cooling due to the Peltier effect compensates for the radiant heat absorbed. If the resistance of the arrangement were so small that the heating due to the Joule effect were negligible, we should have—

$$w=\pi i,$$

where w is the energy in ergs absorbed per second, π the Peltier coefficient, and i the current. But the resistance is never negligible, and the heating due to it being i^2r ergs per second,

$$w=\pi i-i^2r,$$

$$=\pi i\left(1-\frac{ir}{\pi}\right).$$

Calling $\frac{\pi}{r}$ the neutral current i_0, for which the Joule heating is just equal to the Peltier cooling ($i_0{}^2r=\pi i_0$), so that the disc would be neither warmed nor cooled by such a current, we have—

$$w=\pi i\left(1-\frac{i}{i_0}\right).$$

In the actual arrangement employed (Fig. 209) there are two similar discs, 1 and 2, each supported by four stout iron and four

[1] H. L. Callendar, *Proc. Phys. Soc. Lond.*, **23**, Part I. December, 1910.

constantan wires, the two discs being thus the junctions of an iron-constantan couple. The discs also form the junction of the single-wire iron and constantan circuit in which the galvanometer G is included. Suppose that the radiation falls on the disc 1, the arrangement will be as shown in the diagram, and the current is adjusted by the resistance R until the two discs remain at the same temperature, as indicated by the current in G being zero. Knowing the current i, as indicated by the milliammeter A, and the Peltier coefficient π, w, the rate of absorption of radiant energy by the disc 1 is known. The neutral current i_0 is determined by a preliminary measurement in which no radiation falls upon either disc.

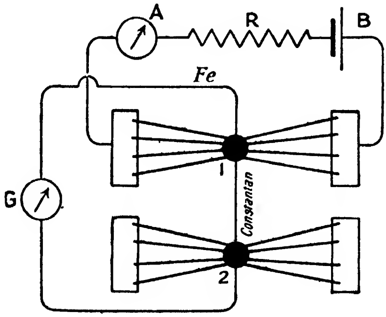


FIG. 209.

Should the whole apparatus become warmed by the radiation falling upon it, as would probably be the case when the sun's radiation is being measured, both discs are affected in the same way, and the error due to this cause is eliminated. Also the radiation may be allowed to fall on the disc 2 instead of 1, for the purpose of making a control measurement.

In a later design the discs are replaced by cups, for the purpose of rendering the absorption more complete ; the electrical arrangements in this case are very similar to those in the disc apparatus.

The "cup" arrangement has also been used as a calorimeter and the Peltier coefficient, for a junction placed in either cup, directly measured. It has also been used for measuring the rate of emission of heat by radioactive substances.

Pyro-electricity.—Another electrical effect due to differences of temperature should be noticed. Certain crystals, especially tourmaline, exhibit electrical charges when heated or cooled. The name *pyro-electricity* is given to this phenomenon. If a crystal of tourmaline be raised in temperature, one end becomes positively and the other negatively charged while the temperature is rising, but during cooling the charges are reversed. This order of the charging takes place whether the crystal be heated or cooled from the atmospheric or from any other temperature. If a crystal be broken up, each part of it exhibits the same properties, and if tourmaline be powdered and spread on a glass plate and warmed, or cooled, the particles gather themselves together in chains, owing to the polar charges, just as iron filings do when magnetised.

It has been thought that only hemi-morphic crystals exhibit pyro-electric properties, but according to Hankel [1] hemi-morphism is not indispensable to the production of pyro-electricity, and it is exhibited by other crystals, provided that their crystallographic axes are unequal; but, in the case of crystals having equal axes, only those which are hemi-morphic are pyro-electric. Boracite, quartz and fluor are among the pyro-electric minerals.

Piezo-electricity.—It was discovered by the brothers Curie [2] that if the crystals which exhibit pyro-electric properties are subjected to compression or tension, opposite charges of electricity appear at the ends of the crystal. Under compression the sign of the charge at either end is the same as would be produced by cooling the crystal, while tension produced charges of the same signs as those due to heating the crystal.

A suitable rectangular block is cut from the crystal, and a sheet of tinfoil laid over each end. The whole is then placed between ebonite blocks, to which the stress is applied. The quadrants of an electrometer being then connected to the tinfoils, the production of charge can be readily investigated.

It was found that the amount of charge produced at each end of a block of tourmaline is proportional to the total force applied to the block and not to the pressure, and that the amounts of positive and negative charge are equal.

The charges produced in this way were used at a later date for the measurement of ionisation current (Chap. XIV) by a compensation method.

The converse effect has been used by Wood and Loomis [3] to produce high frequency oscillations in quartz. Tinfoil sheets cover the opposite faces of a slab of quartz, and the application of an alternating p.d. of 50,000 volts at frequency 300,000 to the foils produces oscillations in the quartz that can be communicated to a liquid.

If the natural frequency of mechanical vibration of the quartz crystal coincides with that of the applied E.M.F., this quartz oscillator is a useful control of the frequency of oscillation in a valve circuit [4] (Chap. XVI). Also, an oscillator of this kind serves as a very useful standard of frequency for oscillatory electric circuits.[5]

[1] Hankel, *Pogg. Ann.*, Bd. **49**, **50**, **53** and **56**.
[2] J. and P. Curie, *Comptes Rendus*, **92**, p. 186. 1881.
[3] R. W. Wood and A. L. Loomis, *Phil. Mag.*, **4**, p. 417. 1927.
[4] G. N. N. Cobbold, *Journ. Inst. El. Eng.*, **66**, p. 855. 1928.
[5] D. W. Dye, *Phys. Soc. Proc.*, **38**, p. 399. 1926.

CHAPTER VIII

ELECTROMAGNETICS

We will now return to the consideration of Ampère's theorem given on p. 48, that an electric current is equivalent to a magnetic shell whose boundary coincides with the current. By a series of experiments, Ampère showed that the magnetic effect at distant points produced by a current, might in all cases be explained by the employment of a magnet or system of magnets, whose polar faces are bounded by the current. Thus a solenoidal current is equivalent to a bar magnet whose ends coincide with the faces of the solenoid, and a wire bent into a circle, when carrying current, is equivalent to a circular magnetic sheet or shell, whose polarity is N on one side and S on the other, the side whose polarity is N depending on the direction of the current. An inspection of Fig. 37 will make it clear which is the N side of the sheet. The following rule will be of assistance.

Imagine the conductor to be placed in the palm of the right hand and the fingers closed upon it, the thumb being outstretched; then if the thumb indicates the direction of the current, the fingers indicate the direction of the magnetic field.

It then follows that if we look upon the N side of a magnetic shell, the current flows in an anti-clockwise direction as seen by the observer.

If the coil has a number of turns, as in the case of a solenoid, the turns being approximately circles, each turn has its equivalent shell, and within the solenoid the N polar face of one shell coincides with the S polar face of the next, and the external effect of these inner shells is zero, but the N at one end of the solenoid, and the S at the other, produce a field similar to that of a circular bar magnet.

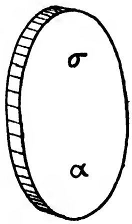


FIG. 210.

Strength of Magnetic Shell.—The magnetic moment per unit area of shell is called the strength of the shell. Thus if σ be the strength of the shell, and a its area, total magnetic moment of shell$=a\sigma$. Further, *we will define the electromagnetic C.G.S. unit of current as one which produces the same magnetic field at external points as a magnetic shell of unit strength whose boundary coincides with the current.*

It is not necessary to define either the thickness of the shell or

the amount of pole per unit area of face, as the magnetic moment of unit area of the shell is equal to the product of these two quantities; the shell is usually considered to be indefinitely thin.

When the current circuit is of very small dimensions, the equivalent shell becomes a small magnet and the magnetic potential and field at any point due to it may be calculated as on p. 15. Thus $V=\frac{M\cos\theta}{r^2}$, and $H=-\frac{dV}{dr}$; but whatever the dimension of the circuit, the method may be extended to give the same quantities. For the circuit ABCD (Fig. 211) may be divided up into a number of meshes by a network of conductors. If now in each mesh a current of strength equal to that in ABCD be considered to flow in the same direction, the side of each mesh not situated at the boundary will have equal and opposite currents flowing in it, and the total currents in the meshes are therefore zero, except at the boundary, where the resultant of the currents in the elements is the current in ABCD. Since each mesh may be replaced by the equivalent element of a magnetic shell of strength equal to the current, the whole shell thus formed is equivalent to the current ABCD.

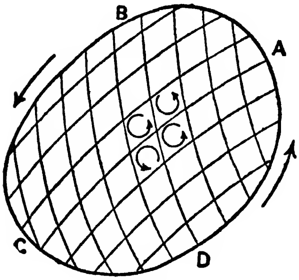


FIG. 211.

The magnetic potential at P (Fig. 212) due to a small element a of the shell is $\frac{a\sigma\cos\theta}{r^2}$, where $a\sigma$ is the magnetic moment of the element. Now if $d\omega$ be the solid angle subtended by a at P, $r^2d\omega$ is the right section of the cone of angle $d\omega$, at distance r from the vertex P, and $\frac{r^2d\omega}{a}=\cos\theta$.

$$\therefore\ d\omega=\frac{a\cos\theta}{r^2}.$$

Hence the potential at P due to this element may be written—

$$dV=\sigma d\omega.$$

And the potential at P due to the whole shell is, $\int dV=\int\sigma d\omega$, but σ being constant—

$$\text{Potential at } P=\sigma\int d\omega=\sigma\Omega,$$

where Ω is the solid angle subtended by the whole shell at P.

It follows that the potential at any point due to a shell depends only upon the strength of the shell and the solid angle subtended by it at the point, and this is independent of any variation in the

shape of the shell, provided that its boundary is fixed. Thus for an infinite plane shell, the potential at neighbouring points is $2\pi \, . \, \sigma$, and for a hemispherical shell the potential at the centre

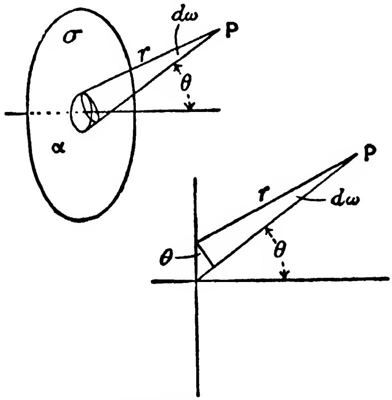


FIG. 212.

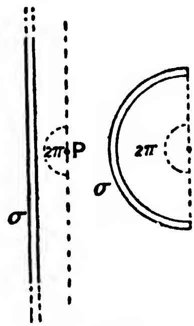


FIG. 213.

is also $2\pi \, . \, \sigma$, since in each case the solid angle subtended at P (Fig. 213) by the shell is 2π.

Circular Current.—If a current of strength i absolute units flow in a circle, we can replace it by a circular magnetic shell of strength $\sigma = i$.

Let AB (Fig. 214) be a side view of the circle; then to find the magnetic potential at a point O on the axis, all that is necessary is to find the solid angle subtended by the circle at O. To do this draw a sphere with centre O, such that the circle lies upon the sphere. Then $\dfrac{\text{area of slice ACB}}{r^2} = \Omega$, where Ω is the solid angle required. Now it may be shown geometrically that the area of a slice of a sphere lying between two parallel planes is equal to the area of the circumscribing cylinder between the planes, and whose axis is perpendicular to these planes. The area of the slice ACB is therefore $2\pi r h$, where $h = r - x$, x being the distance of O from the plane of the circle.

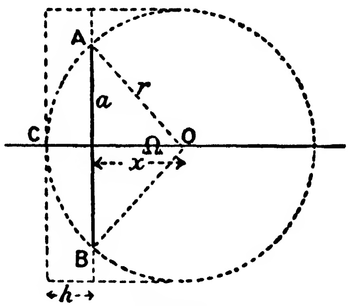


FIG. 214.

$$\therefore \ \Omega = \frac{2\pi r h}{r^2} = \frac{2\pi h}{r} = \frac{2\pi(r - x)}{r}$$

$$= 2\pi\left(1 - \frac{x}{r}\right).$$

But, magnetic potential $V=\sigma\Omega=i\Omega$;

$$\therefore\ V=2\pi i\left(1-\frac{x}{r}\right)$$

Noting that $r^2=x^2+a^2$, we have—

$$V=2\pi i\left\{1-\frac{x}{(x^2+a^2)^{\frac{1}{2}}}\right\}.$$

By symmetry we see that the magnetic field due to the circular current is directed along the axis, and its value is therefore $-\frac{dV}{dx}$ (see p. 13).

$$\therefore\ H=-\frac{dV}{dx}=-2\pi i\frac{d}{dx}\{1-x(x^2+a^2)^{-\frac{1}{2}}\}$$

$$=2\pi i\{-\tfrac{1}{2}\,.\,x\,.\,2x(x^2+a^2)^{-\frac{3}{2}}+(x^2+a^2)^{-\frac{1}{2}}\}$$

$$=\frac{2\pi a^2 i}{(x^2+a^2)^{\frac{3}{2}}}.$$

For a point at the centre of the circle, $x=0$, and then $H=\frac{2\pi i}{a}$; which is in accordance with the result derived from the law on p. 49.

Solenoidal Current.—When the current is flowing in a cylindrical sheet, its direction being everywhere perpendicular to the axis, it is said to be solenoidal, and the strength of magnetic field inside it may be found from the above relation. This condition is very nearly fulfilled by a current flowing in a wire closely wound upon a cylinder, when the thickness of the wire is small compared with the radius of the cylinder.

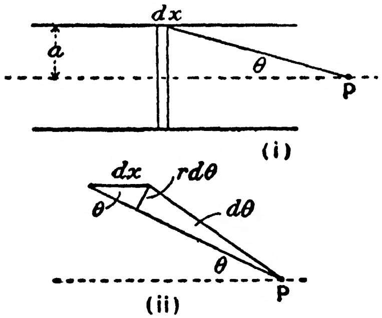


FIG. 215.

If i be the current per unit length of the solenoid, idx is that in a thin section of length dx (Fig. 215 (i)). The strength of field at P due to this is $\frac{2\pi a^2 i dx}{(x^2+a^2)^{\frac{3}{2}}}$, where x is the distance of the plane of the circle from P. But, from the enlarged diagram (Fig. 215 (ii)) we see that $rd\theta=dx\,.\,\sin\theta$,

$$\therefore\ dx=\frac{r\,.\,d\theta}{\sin\theta}.$$

$$\text{And field due to section}=\frac{2\pi a^2 i\,.\,rd\theta}{(x^2+a^2)^{\frac{3}{2}}\sin\theta}.$$

Remembering that $\frac{a}{r}=\sin\theta$, and $x^2+a^2=r^2$, we write the expression in the form—

$$\frac{2\pi i a^2 r \,.\, d\theta}{r^3 \sin\theta}=\frac{2\pi i a^2 \,.\, d\theta}{r^2 \sin\theta}$$

$$=2\pi i \,.\, \sin\theta \,.\, d\theta.$$

And for the whole solenoid, $H=2\pi i\int_{\theta_1}^{\theta_2}\sin\theta \,.\, d\theta$

$$=2\pi i\Big[\cos\theta\Big]_{\theta_2}^{\theta_1}$$

where θ_1 and θ_2 are the values of θ at the ends of the solenoid.

If the solenoid consists of wire of n turns per centimetre length, the current in each turn being i,

$$H=2\pi n i\Big[\cos\theta\Big]_{\theta_2}^{\theta_1}=2\pi n i[\cos\theta_1-\cos\theta_2].$$

When the length of the solenoid is infinite, $\theta_1=0$, and $\theta_2=\pi$, and therefore—

$$H=4\pi n i.$$

Work done in carrying a Magnetic Pole round a Current.—Remembering that the difference in magnetic potential between two points is the work done in carrying a unit N pole from one point to the other, and that it is independent of the path along which the pole is carried, we may prove one of the most important laws connecting electric and magnetic quantities.

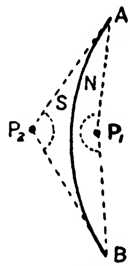


FIG. 216.

Consider two points P_1 and P_2 very close to, but on opposite sides of, the magnetic shell AB (Fig. 216) of which the N polar face is towards P_1. The magnetic potential at P_1 due to the shell is $+\sigma\times$(solid angle AP_1B), the solid angle AP_1B being shown by the dotted arc in the figure. Similarly the magnetic potential at P_2 is $-\sigma\times$(solid angle AP_2B)

$\therefore$ Difference of potential between P_1 and P_2

$$=\{\sigma(\text{solid angle } AP_1B)\}-\{-\sigma(\text{solid angle } AP_2B)\}$$
$$=\sigma(\text{solid angle } AP_1B+\text{solid angle } AP_2B).$$

As the points P_1 and P_2 approach each other, the sum of the solid angles AP_1B and AP_2B becomes more and more nearly equal to the solid angle subtended by the whole of space surrounding a point, that is to 4π, and since the magnetic shell may be considered to be indefinitely thin, the points P_1 and P_2 may approach each other until their distance apart is infinitesimal,

and still they are on opposite sides of the shell. Hence the difference of magnetic potential between two points P_1 and P_2 on opposite sides of, and very close to, a magnetic shell is $4\pi\sigma$, which is then the work done in carrying unit pole from P_1 to P_2 by any path, provided that the path does not intersect the shell. On passing through the shell from P_2 to P_1, the direction of the force on the unit pole is reversed, and if the work were calculated it would be found to be equal and opposite to that for the external path from P_1 to P_2. It is not necessary to perform this calculation, because the potential at a point such as P_1, due to any distribution of magnetisation, can only have one value, so that the total work for a closed path is zero ; otherwise useful work might be done by allowing a pole to circulate round a closed path, without any corresponding loss of energy in the system, and this is contrary to experience.

If, however, the shell be replaced by its equivalent current ($\sigma=i$) flowing round the boundary of the shell AB, the work for the external path from P_1 to P_2 is $4\pi\sigma$, or $4\pi i$, as in the case of the shell, but now the work required to complete the path in going from P_2 to P_1 may be made as small as we please by taking P_1 and P_2 sufficiently close together, there being in this case no magnetic material to traverse. A closed path such as we have described is necessarily linked once with the current, and thus *the work done in carrying a unit pole round a closed path linked once with a current i is $4\pi i$.* Thus the magnetic potential at any point in the neighbourhood of a current may be considered to have a number of potentials whose values differ by multiples of $4\pi i$. The potential due to a current is therefore multi-valued. There is, in this, no contradiction to the principle of the conservation of energy, for the current is not a statical phenomenon ; it has to be maintained by the continuous expenditure of energy in the battery. When the magnetic pole is carried round the circuit its field cuts the circuit during the process and produces current which, if in the opposite direction to the principal current, will cause a temporary lessening of the rate of expenditure of energy in the battery ; if in the other direction, an increase in its rate. In either case we can trace the source, or the mode of disposal, of the energy corresponding to the work done in carrying the pole round the path linked with the current, to the change in the amount of chemical action occurring in the battery. After the completion of the path, the circuit is not in the same condition as at the start.

Line Integral of Magnetic Field.—The work done in carrying a unit pole along any path from one point to another is called the line integral of the field between the points. If the strength of field at any point of the path be H, and θ its inclination to the

poles m_1 and m_2 situated in a magnetic medium of susceptibility κ is $\frac{m_1 m_2}{(1+4\pi\kappa)r^2}$ the factor $(1+4\pi\kappa)$ involving 4π due to the effect of the magnetisation of the medium. As before, it is convenient to substitute a symbol for $1+4\pi\kappa$, and the one chosen is μ. It is called the *magnetic permeability* of the medium.

Thus, $$\mu=1+4\pi\kappa.$$

and, $$\text{force between poles}=\frac{m_1 m_2}{\mu r^2}.$$

It follows, from above, that $B=\mu H$.

The permeability μ is not, like the dielectric constant k, a constant for any one material. In the case of iron, nickel and cobalt the value of μ is much greater than for other substances and for any one specimen its value varies between wide limits.

Demagnetisation.—It must be understood that H in the above expression is the actual field producing magnetisation within the material, and that if there are free poles upon the specimen they will always produce a field which is in opposition to, and must be subtracted from, the original field, within the material, in order to obtain the resultant magnetising effect. Thus, for a magnet NS (Fig. 256) each pole produces its own radial field, the resultant being the ordinary field due to a pair of poles. At the middle of the magnet this field is opposed to the magnetising field H, and therefore exerts a demagnetising effect upon the bar. It is for the purpose of removing the free poles that produce this demagnetising effect, that permanent magnets are usually provided with soft-iron keepers, the keeper producing poles equal and opposite to those of the magnet, and being very nearly coincident in position with them, these poles produce a field equal and opposite to the demagnetising field.

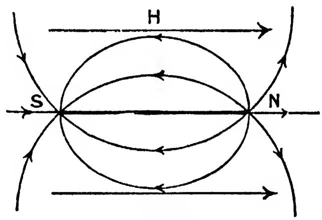


FIG. 256.

Whatever the form of the magnet, the demagnetising field is proportional to the strength of the pole to which it is due, and this in turn is proportional to the intensity of magnetisation, so that the demagnetising field is equal to NI, where N is a constant depending on the geometrical form of the magnetised body.

If then H′ is the magnetising field when the body is absent, and H that actually existing in the interior of the body—

$$H=H'-NI.$$

N may be calculated in a number of simple cases when the

If the thickness of the wire be taken into account, so that its form is that of a solid cylinder, the field outside it is $\frac{2i}{r}$, at a distance r from the axis, but that inside it will be different. For let r_1 be the distance from the axis of a point inside the wire, and let i_1 be the current density, that is the current flowing through unit area of cross-section of the wire. This is uniform over the whole section when the current is steady, and therefore the total current within the cylinder of radius r_1 is $\pi r_1{}^2 i_1$. The field at distance r_1 due to this current is

$$\frac{2(\pi r_1{}^2 i_1)}{r_1}=2\pi r_1 i_1=\frac{2r_1 i}{r^2},$$

where i is the whole current $\pi r^2 i_1$.

This is the actual strength of field, since the current in the cylindrical shell lying outside the point does not produce any field within it, the circular path inside not enclosing any of this current.

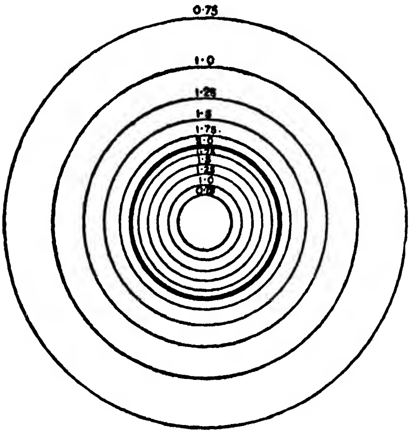


FIG. 219.

Thus the field due to a current in a cylinder is greatest at the surface of the cylinder, its value being there $\frac{2i}{r}$, and it falls off as we pass either outwards or inwards, being zero at the axis. The distribution of the magnetic lines of force is shown in Fig. 219, the values of H being marked upon the circles when that at the surface of the conductor is taken to be 2·0 C.G.S. units.

Magnetic Field due to Solenoid.—For a solenoid in the form of a ring, frequently called an endless solenoid, the line integral of the field round the axis of the solenoid (Fig. 220) is $2\pi r$H. If, then, there are n turns of wire per centimetre length of solenoid, there are in all $2\pi rn$ turns, and the circular path of radius r is linked $2\pi rn$ times with the current. If, then, i be the current in each turn, the effective current linked with the path is $2\pi rni$, and it follows from the law given on p. 230, that

$$2\pi r\mathrm{H}=4\pi(2\pi rni),$$
$$\therefore\ \mathrm{H}=4\pi ni.$$

It will be noticed that r will vary slightly according to whether the path is near the inner or the outer surface of the solenoid, and therefore the field is not quite uniform; but when the thickness of the solenoid is small compared with its radius, this departure from uniformity of field is negligible, and if r be infinite, the solenoid is a straight one, and the field inside it is uniform, its value being $4\pi ni$. This is in agreement with the result obtained on p. 229.

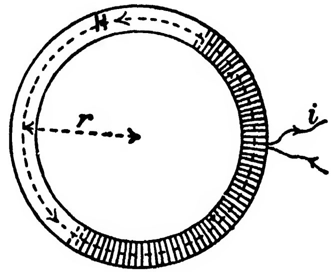


FIG. 220.

Magnetic Permeability.—There is a close mathematical analogy between magnetic fields and statical electric fields, due to the similarity in the laws of force between magnetic poles and that between electric charges. In the magnetic, as in the electrical case, the force depends upon the medium in which the poles are situated. It is convenient to determine the unit of magnetic pole from the force between poles situated *in vacuo*, and this is practically the same as for air; but there are many media for which the force between the poles differs greatly from that between the same poles situated in air or *in vacuo*. We must, therefore, rewrite our force equation in the form—

$$\text{Force}=\frac{m_1 m_2}{\mu r^2},$$

where μ is a quantity depending upon the medium in which the poles are situated. It is called the *magnetic permeability* of the medium, for a reason to be given later.

On p. 3 we saw that the strength of field due to a pole of strength m at a distance r is $\frac{m}{r^2}$; but we now see that when the medium filling the space has permeability μ, field strength $H=\frac{m}{\mu r^2}$.

Again, the magnetic potential due to any distribution of poles is changed from V to $\frac{V}{\mu}$ when the medium is changed from air to one of permeability μ; in fact, the magnetic equations are modified by the quantity μ in exactly the same way as we saw in Chapter IV the equations for a statical electric field to be modified by the dielectric constant k; but this difference should be noted, that whereas k is constant for any given medium, μ is by no means constant: its complex variations will be studied

in Chapter IX. Still μ has a definite value under any given circumstances, defined by the equation $F=\frac{m_1m_2}{\mu r^2}$, although this value may vary at different times and under different conditions.

Magnetic Induction.—The quantity $\mu H \cos \theta . ds$ is defined as the *magnetic flux* over the surface ds, where θ is the angle between H and the normal to ds; and *Gauss's Law* in the magnetic case may be proved exactly as on p. 123 for the electrical case. Thus the total magnetic flux over a closed surface is equal to 4π times the amount of pole within it; and

$$\int \mu H \cos \theta . ds = 4\pi \Sigma m.$$

It follows as on p. 126, that the strength of magnetic field due to a plane polar sheet is $2\pi\sigma_0$, where σ_0 is the amount of pole per unit area of the sheet.

We give a special name to the quantity μH: it is, the *Magnetic Induction* (B) and is analogous to ϕ in the electrical case (p. 118), thus—

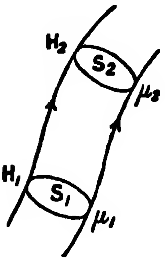


FIG. 221.

$$\phi = kE, \quad B = \mu H,$$

and the magnetic field may be mapped out by means of tubes of induction, whose characteristic property is that BS is constant for any tube. Thus in Fig. 221, if H_1, H_2 and μ_1, μ_2 are the values of H and μ at the sections of the tube of induction having areas S_1 and S_2, the flux over the sides of the tube is zero, their direction being everywhere that of the field, we have from Gauss's law when there is no pole within the tube—

$$\mu_1 H_1 S_1 = \mu_2 H_2 S_2$$

or, if μ is constant, $$B_1 S_1 = B_2 S_2.$$

Boundary Conditions.—Following the analogy we see, as on p. 138, that the boundary conditions that must be satisfied at the surface of separation of two media of different magnetic permeabilities are—

(i) The tangential components of the field are the same in both media;

i.e. $$H_1 = H_2$$

and (ii) The normal components of the magnetic induction are the same in the two media;

i.e. $$B_1 = B_2, \text{ or, } \mu_1 H_1 = \mu_2 H_2.$$

Thus for a line of induction which crosses the boundary we have (Fig. 222)—

from (i) $$H_1 \sin \theta_1 = H_2 \sin \theta_2,$$

and from (ii) $$\mu_1 H_1 \cos \theta_1 = \mu_2 H_2 \cos \theta_2.$$

Dividing one equation by the other, we have—

$$\frac{\tan\theta_1}{\tan\theta_2}=\frac{\mu_1}{\mu_2}.$$

The problem of a sphere of permeability μ_2 situated in a medium of permeability μ_1, the original field H being uniform, is exactly analogous to that for the dielectric sphere in an electrical field E (pp. 139 to 142), and the argument may be repeated, replacing E by H, k_1 and k_2 by μ_1 and μ_2. The resultant field H_2 inside the sphere is thus

$$H_2=\frac{3\mu_1}{\mu_2+2\mu_1}H \qquad \text{(p. 141).}$$

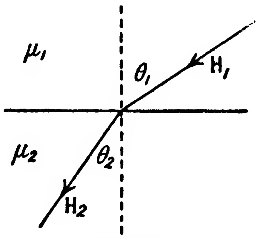


FIG. 222.

Fig. 143 may illustrate the case in which $\mu_2>\mu_1$, as in the case of the magnetic metals, iron, nickel and cobalt, situated in air; these substances are said to be ferromagnetic. In Fig. 144, $\mu_1>\mu_2$, a condition which is fulfilled for some substances in air, in which case the substance is said to be diamagnetic. Bismuth, for which $\mu=0.99997$, is one of the most strongly diamagnetic substances known.

Magnetic Shielding.—The tendency of the magnetic tubes of induction to concentrate upon places of high permeability explains the use of hollow iron spheres and cylinders to reduce the magnetic field in the spaces within them. It is sometimes desirable to protect a suspended-needle galvanometer from magnetic disturbances, and although this can never be completely effected, the disturbing field may be very much reduced by surrounding the instrument by massive iron shields. The calculation of the change in field produced is beyond the scope of our present work, but the results for a sphere and a cylinder are of use. They have been given by du Bois.[1]

The field inside a hollow sphere is—

$$\frac{H}{1+\frac{2}{9}(\mu-2)\left(1-\frac{r^3}{R^3}\right)},$$

where H is the external field, μ the permeability, and r and R the internal and external radii. For a cylinder with axis at right angles to the field it is—

$$\frac{H}{1+\frac{1}{4}(\mu-2)\left(1-\frac{r^2}{R^2}\right)}.$$

[1] H. du Bois, *Electrician*, 40, p. 317. 1898.

The nickel-iron alloy known as mumetal has a permeability which varies from 10,000 to 100,000. This high permeability renders it useful for magnetic shielding.

Force on Magnetic Body in Magnetic Field.—We saw on p. 137 that a body whose dielectric constant is greater than that of the surrounding medium, situated in an electric field that is not uniform, tends to move towards the stronger parts of the field, and the same consideration would lead us to a like conclusion in the case of a paramagnetic or ferromagnetic body. Since the force on small bodies has been used for measuring their magnetic properties, we will calculate the force on such bodies.

Let the body consist of two poles of strength m, the magnetic potential where the poles are situated being V_n and V_s.

$$\text{Then potential energy of body} = m(V_n - V_s),$$

being the work done in bringing the body from infinity to the point, and if the distance between the poles be ds—

$$\text{potential energy} = mds \,.\, \frac{V_n - V_s}{ds}$$
$$= M\frac{dV}{ds},$$

when ds becomes sufficiently small, and M is the magnetic moment of the body. Further, $\frac{dV}{ds} = -H$, and $M = vI$, where v is the volume of the body and I the intensity of magnetisation (see p. 270).

$$\therefore \text{potential energy} = -vIH.$$

Now the work done during a small displacement of the body is the difference in the potential energy before and after the displacement, and it is also equal to the product of the force F in the direction of the displacement and ds the amount of displacement.

$$\therefore Fds = -d(-vIH),$$
$$\text{or,}\quad F = vI\frac{dH}{ds}.$$

For all feebly magnetised bodies, including those which are paramagnetic or diamagnetic the magnetic susceptibility κ is constant. Now $I = \kappa H$, $\therefore F = v\kappa\frac{HdH}{ds} = \frac{v}{2}\kappa\frac{dH^2}{dx}$.

In the case of a sphere of susceptibility κ, placed in a field of strength H, we shall see on p. 272 that—

$$I=\frac{\kappa H}{1+\frac{4}{3}\pi\kappa},$$

$$\therefore F=\frac{v\kappa}{1+\frac{4}{3}\pi\kappa}\cdot\frac{HdH}{ds},$$

$$=\frac{v}{2}\cdot\frac{\kappa}{1+\frac{4}{3}\pi\kappa}\cdot\frac{dH^2}{ds}.$$

Hence when the field is uniform H^2 is constant, and there is no force on the body, and further, the direction of the greatest value of F is that in which H^2 varies most rapidly; again when κ is negative, or is less than that of the surrounding medium, the direction of F is reversed. Since paramagnetic bodies tend to move from the weaker to the stronger parts of the field, diamagnetic bodies tend to move towards the weaker parts of the field.

Equivalence of Current and Magnetic Shell in any Medium.—The work done in carrying a unit magnetic pole round a closed path linked with a current is independent of the presence of any distribution of magnets there may be, since the work done in traversing the closed path, and due to any magnets in the neighbourhood, we have seen to be zero (p. 230). It follows that if there are magnets or magnetic material in the neighbourhood of the current, they will not change the amount of work done in carrying a unit magnetic pole round the closed path, which is therefore always $4\pi i$.

If then, the whole of space is filled with a medium of permeability μ, differing from unity, the magnetic field is everywhere the same as when the permeability was unity, and the work done by our unit pole in its circuital path is still $4\pi i$. If the space be partly filled with magnetic material the work is still $4\pi i$, but owing to the presence of free poles at the boundary of the magnetic material, the field will be increased at some points and diminished at others, a fact which will be seen in dealing with the demagnetising effect in the interior of a mass of iron in a magnetic field (p. 272).

We must modify our conception of the equivalent magnetic shell for a given current circuit to bring it into accordance with these ideas. For we see that on filling space with a material of permeability μ, the field everywhere due to the current is unchanged, but that due to the magnetic shell is reduced to $\frac{1}{\mu}$ of its previous value (p. 233). Hence if the shell is still to be equivalent to the current we must increase its strength μ times, and we may then say that *in a medium whose permeability is everywhere μ, the current is equivalent to a magnetic shell of strength σ, where $\sigma=\mu i$.*

Force on Current in Magnetic Field.—The equivalence of a current and a magnetic shell leads us to the conclusion that a conductor in which an electric current is flowing will experience a force when situated in a magnetic field; in fact, it was by a series of experiments in which the forces on circuits carrying current were produced by magnets, that Ampère established the equivalence of a current and a magnetic shell. The direction and magnitude of the force may be found as follows.

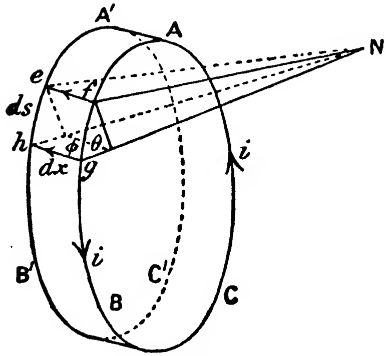


FIG. 223.

Consider the circuit ABC, in which current i is flowing, to be displaced always parallel to itself through distance dx, so that its new position is A′B′C′ (Fig. 223). Owing to the presence of a N pole of strength m, situated at N, work is done when the displacement occurs, and the potential energy of the system consisting of the pole and the current is changed by an amount equal to the work done in displacing the circuit from the first to the second position.

Let Fds be the force on element ds of the conductor, acting in the direction of the displacement, F being the force per unit length at this part of the circuit. Then Fds . dx is the work done on the element ds during the displacement. And for the whole circuit

$$\text{work done during displacement} = \Sigma F ds \,.\, dx.$$

The area swept out by the element ds is that of the figure $efgh = ds \,.\, dx \,.\, \sin\phi$; and the solid angle subtended by this at the point N is—

$$\frac{(\text{area } efgh) \sin\theta}{r^2} = \frac{ds \,.\, dx \,.\, \sin\phi \,.\, \sin\theta}{r^2},$$

where θ is the angle between the line joining N to $efgh$, and the plane of $efgh$; therefore the solid angle subtended at N by the whole curved surface ABCC′B′A′ is—

$$\Sigma\frac{ds \,.\, dx \,.\, \sin\phi \sin\theta}{r^2}.$$

Now if the solid angles subtended by the circuits ABC and A′B′C′ at N be respectively Ω and Ω',

$$\Omega - \Omega' = \text{solid angle subtended by ABCC′B′A′}$$
$$= \Sigma\frac{ds \,.\, dx \,.\, \sin\phi \sin\theta}{r^2}.$$

But change of potential at N produced by the displacement of the circuit (p. 226)

$$=i\Omega-i\Omega'=i\Sigma\frac{ds\,.\,dx\,.\,\sin\phi\sin\theta}{r^2}.$$

The change in potential measures the difference in the amounts of work required to bring unit pole from infinity to N, when the circuit is at ABC and A′B′C′ respectively, and is therefore the change in potential energy of the system for the given displacement when the pole has unit strength. But since the pole has strength m, this change in potential energy is—

$$=mi\Sigma\frac{ds\,.\,dx\,.\,\sin\phi\sin\theta}{r^2},$$

$$\therefore\ \Sigma\mathrm{F}ds\,.\,dx=mi\Sigma\frac{ds\,.\,dx\,.\,\sin\phi\sin\theta}{r^2}.$$

And for this equation to be satisfied—

$$\mathrm{F}=\frac{mi\sin\phi\sin\theta}{r^2}.$$

The greatest value of this for a given value of θ occurs when $\phi=90°$, that is, the force is greatest in a direction at right angles to that of the current, in which case $\mathrm{F}=\frac{mi\sin\theta}{r^2}$, and for an element ds—

$$\text{Force}=\frac{mi\,.\,ds\sin\theta}{r^2},$$

and further, for a given displacement in any direction, the work done, and therefore the force on the element, is greatest when the solid angle subtended by the circuit is changed most for that displacement, and this is greatest when it is at right angles to r, which is the direction of the magnetic field due to N. The resultant force on the element is therefore always at right angles to the magnetic field, and we have seen that it is at right angles to the element ds, and hence it is at right angles to the plane containing the element of the current and the direction of the magnetic field.

We see from Fig. 223 that in this case θ is the angle between the current and the field, and therefore the force per unit length of conductor is $\frac{mi\sin\theta}{r^2}$. But $\frac{m}{\mu r^2}$ is the strength of magnetic field H due to the pole, and $\mu\frac{m}{\mu r^2}\left(=\frac{m}{r^2}\right)$ is the induction B due to it.

$\therefore$ Force per unit length of conductor$=$Bi sin θ, and is at right angles to B and to i. The force is H i sin θ when $\mu=1$.

It will be seen in Fig. 224 that the directions of the quantities *i*, H and F are related to each other as in Fig. 223, which may be remembered by Fleming's Left Hand Rule. If the thumb, fore and middle fingers of the left hand are extended so that they are mutually at right angles, and the m*I*ddle finger point in the direction of the current (I), the *F*ore finger in the direction of the magnetic *F*ield, then the thu*M*b points in the direction of *M*otion when the circuit moves owing to the action of the field.

Another useful way of expressing this fact is that if we look along the magnetic field then an anti-clockwise rotation brings the direction of the current into coincidence with that of the motion, that is the "cause" into the direction of the "effect." This latter

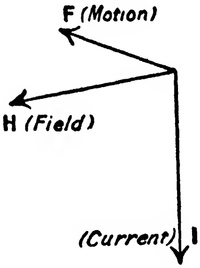


FIG. 224.

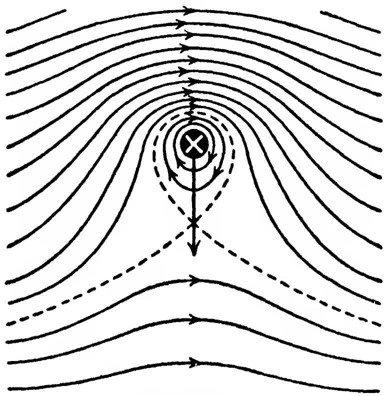


FIG. 225.

rule has the advantage of being the same as that for the induced E.M.F. (p. 252).

The same conclusion regarding the direction of the force experienced by a current situated in a magnetic field may be reached by considering the magnetic lines of force of the resultant field. Those of the original magnetic field are parallel straight lines and those due to the linear current are circles. The magnetic lines for the current and field combined are shown in Fig. 225, when the current flows downwards through the plane of the paper. The lateral pressure between the tubes of force above the wire where they are crowded together is greater than below it, and the result will be that the wire experiences a force which is directed downwards in the diagram; it will be seen that this is the direction previously found for it. The dotted line is the boundary separating those lines which pass on one side of the current from those on the other side.

Suspended Coil.—We can now explain the use of a suspended coil for galvanometric purposes; for a rectangular coil carrying

current experiences a couple when suspended in a magnetic field. The rectangular coil *abcd* (Fig. 226 (i)) may be considered to be the suspended coil of a galvanometer. Then the force per unit length of *ab* and *dc* is Hi, and the total force Hi(*ab*), the direction of the forces being shown in the plan (Fig. 226 (ii)). These give rise to a couple Hi(*ab*)(*ed*) tending to twist the coil into the position in which its normal has the same direction as the field.

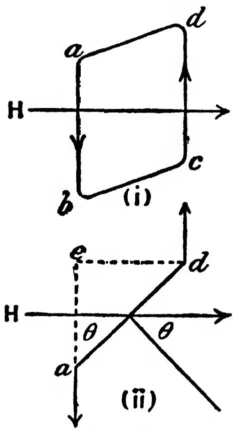


FIG. 226.

$$\text{Couple}=\mathrm{H}i(ab)(ed)=\mathrm{H}i(ab)(ad)\sin\theta$$
$$=\mathrm{H}i\mathrm{A}\sin\theta,$$

where $\mathrm{A}=ab\times ad$, the area of the coil. The sides *ad* and *bc* do not contribute anything to the couple, since the forces on them are vertical, that on *ad* being vertically upwards and that on *bc* vertically downwards.

The couple HiA sin θ might have been derived directly by replacing the circuit by its equivalent magnet shell, whose magnetic moment is iA, and is in the direction of the normal to the coil. The couple on this is HiA sin θ.

If the uniform field H be due to a permanent magnet, and the coil be suspended by a metallic wire which exerts a controlling couple $c\theta'$, where θ' is the angle between the plane of the deflected coil and the field, equilibrium is attained when—

$$c\theta'=\mathrm{H}i\mathrm{A}\cos\theta',$$

or,

$$i=\frac{c}{\mathrm{HA}}\cdot\frac{\theta'}{\cos\theta'}.$$

The instrument is very much simplified by employing a radial field, in which case the vertical sides of the coil seen at *a* and *d* in Fig. 227, experience forces Hil always at right angles to the plane of the coil, where l is the length of the vertical side. The deflecting couple is Hil(*ad*)=HiA.

The coil therefore comes to rest when—

$$c\theta=\mathrm{H}i\mathrm{A},$$

or,

$$i=\frac{c}{\mathrm{HA}}\theta,$$

and the current is directly proportional to the deflection.

The radial field has the advantage that the couple due to the current does not depend on the position of the coil, whereas in uniform field it varies as the sine of the angle between the field and the normal to the coil.

Effect of Current on Current.—From Ampère's law of the equivalence of a current circuit to a magnetic shell, we should expect that forces would exist between two circuits carrying current. Such effects may easily be produced, and their magnitudes may be calculated from the forces between the equivalent shells. Thus, for two circular currents mutually at right angles (Fig. 228) where AB is a large circle and CD a small one, we

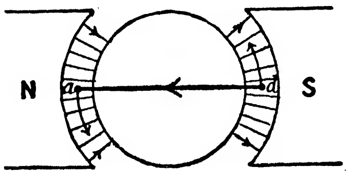


FIG. 227.

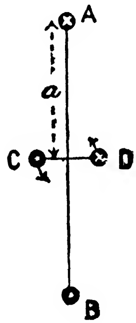


FIG. 228.

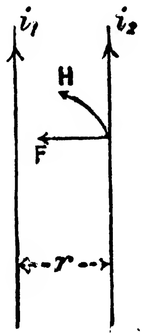


FIG. 229.

have seen (p. 228) that the field at the centre of AB is $\frac{2\pi i_1}{a}$, where a is the radius and i_1 the current; the magnetic moment of the small coil is αi_2 where α is its area and i_2 the current in it. Hence CD will experience a couple $\frac{2\pi i_1 i_2 \alpha}{a}$ tending to twist its plane into that of AB when the two are at right angles, or $\frac{2\pi \alpha i_1 i_2}{a} \sin\theta$ when the planes of the two coils are inclined to each other at an angle θ.

Again, the long straight current i_1 (Fig. 229) produces a magnetic field $\frac{2i_1}{r}$ at a distance r from it, and a second straight current i_2 parallel to the first will experience a force Hi_2 or $\frac{2i_1 i_2}{r}$ per unit length, and it will be seen that when the currents are in the same direction the force urges i_2 towards i_1; when the currents are in opposite directions the force drives i_2 away from i_1. In one case the force per unit length of i_1 is equal and opposite to that in the other case, and we see that currents in the same direction attract each other; those in opposite directions repel each other.

There is a useful method of drawing the magnetic lines of force due to two parallel straight currents. If A and B (Fig. 230 (i))

are the sections of the wires carrying the currents, and *ag* and *cb* are parts of circles with centre at A, they may be considered as neighbouring lines of force due to the current A. Similarly *ca* and *bg* are two lines of force due to B. Imagine the figure *acbg* to be so small that it may be looked upon as a parallelogram, shown in enlarged view at (ii). Draw the radii A*ce* and B*gf*. Then *cfg* and *ceg* are right angles and *cfeg* is therefore a cyclic quadrilateral and $\widehat{ace}=\widehat{agf}$.

$$\text{Then} \qquad ac=\frac{ce}{\cos ace} \quad \text{and} \quad ag=\frac{fg}{\cos agf}$$

$$\therefore \frac{ac}{ag}=\frac{ec}{fg}.$$

If now $Ac=r$, $ce=dr$, $Bg=r'$, $fg=dr'$, then $\frac{ac}{ag}=\frac{dr}{dr'}$. Now *ac* and *ag* are in the direction of the fields due to B and A respectively, and if, in addition, the lines of force are so drawn that $\frac{dr}{dr'}=\frac{\text{field due to B}}{\text{field due to A}}$, then they are in length proportional to the field strengths and *agbc* is the parallelogram of forces for the fields. If the currents A and B are in the same direction, the

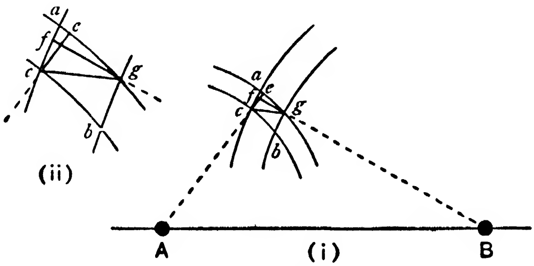


FIG. 230.

diagonal *cg* gives the resultant field, but if the currents are in opposite directions the diagonal *ab* must be used.

For the lines to be spaced properly the number per unit area must be proportional to the field strength. If two planes 1 cm. apart and perpendicular to the field, that is parallel to the plane of the diagram, are considered, the number of lines crossing one radial square centimetre is inversely proportional to the distance *dr* between the lines in the diagram, that is field strength$=\frac{\text{const.}}{dr}$. But the field strength varies inversely as r (p. 231), that is field strength$=\frac{\text{const.}}{r}$. From these it follows that $\frac{dr}{r}$ is constant,

which affords an approximate method of drawing the lines. For if two straight lines EF and EG (Fig. 231) are drawn and

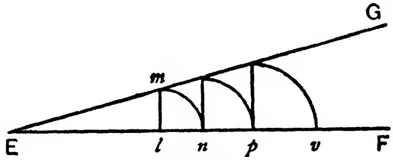


FIG. 231.

El taken as the radius of one line of force, and lm drawn perpendicular to EF, $\frac{ml}{\mathrm{E}l}$=tan GEF. With l as centre draw arc mn and take ln as dr. Repeat the process finding the points p, v, etc. and if these are successive values of dr, then El, En, Ep, Ev, etc. are successive values of r, and $\frac{dr}{r}$=const. In this way the circles in Fig. 232 have been drawn, and in the upper half of the figure

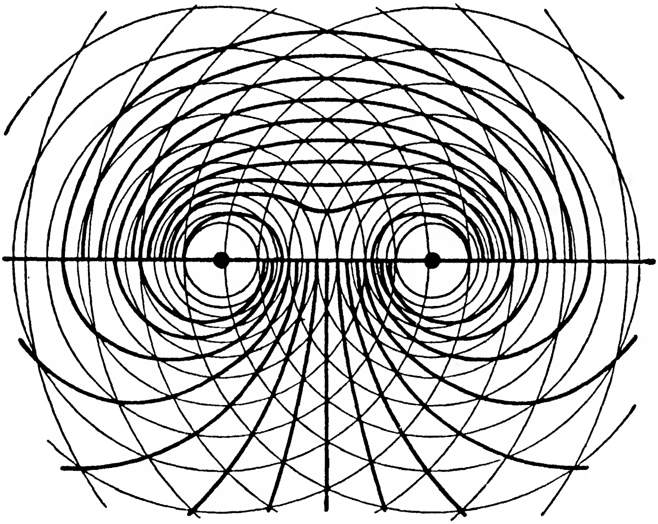


FIG. 232.

the lines of force have been drawn for the currents A and B in the same direction and in the lower half, those for currents in opposite directions. It will be seen that when the currents are in the same direction, the two wires are surrounded by lines or tubes of force, which, by their contraction, would urge the wires together. When the currents are in opposite directions there are no tubes of force surrounding both wires, and since they are more crowded in the space between the wires than in that

outside, the lateral pressures of the tubes will urge the wires apart.

In the same way Fig. 225 may be drawn. The lines for the current are drawn, as above, and a set of equally spaced straight lines to represent the uniform field. The diagonals of the little parallelograms are then drawn.

Coaxial Coils.—(i) For two circular coaxial coils of very nearly the same radius, situated a small distance apart, the force on each unit of length of either coil is $\frac{2i_1i_2}{ab}$ (Fig. 233) in the direction

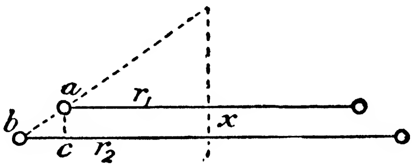


FIG. 233.

ab. The component of this, normal to the axis, taken all round the coils, will, by symmetry, vanish, but the component parallel to the axis is—

$$\frac{2i_1i_2}{ab}\cdot\frac{ac}{ab}=\frac{2i_1i_2\,.\,x}{(ab)^2} \text{ for unit length,}$$

and for the whole circle, since total length is $2\pi r_1$, r_1 being very nearly equal to r_2—

$$\text{Force}=\frac{2i_1i_2\,.\,x\,.\,2\pi r_1}{(r_2-r_1)^2+x^2}$$

$$=\frac{4\pi i_1i_2r_1\,.\,x}{(r_2-r_1)^2+x^2}.$$

This is zero when $x=0$, *i.e.* when the coils are in the same plane, and its maximum occurs when $\frac{x}{A^2+x^2}$ is a maximum, putting A^2 in place of $(r_2-r_1)^2$.

Now, $\frac{d}{dx}\left(\frac{x}{A^2+x^2}\right)=\frac{(A^2+x^2)-2x^2}{(A^2+x^2)^2}=\frac{A^2-x^2}{(A^2+x^2)^2}.$

Putting this equal to zero we have $A^2=x^2$. On obtaining $\frac{d^2}{dx^2}\left(\frac{x}{A^2+x^2}\right)$ and substituting A^2 for x^2 the result is negative, and therefore $A^2=x^2$ or $(r_2-r_1)^2=x^2$ corresponds to a maximum. The force between the coils is therefore a maximum when $x=r_2-r_1$, and its value is then—

$$\frac{2\pi i_1i_2r_1}{r_2-r_1}.$$

When the coil r_1 is so small that the variation of the field over its surface, due to the coil r_2, is negligible, let H be the field due to r_2 (Fig. 234). Then if the magnetic shell due to r_1 have thickness dx, and pole strength m per unit of area—

$$\text{force on under face}=\mathrm{H}\,.\,m\,.\,\pi r_1^2,$$

$$\text{field at upper face}=\mathrm{H}+\frac{d\mathrm{H}}{dx}\,.\,dx,$$

$$\text{and, force on upper face}=\left(\mathrm{H}+\frac{d\mathrm{H}}{dx}\,.\,dx\right)m\,.\,\pi r_1^2.$$

The resultant force on the small coil is the difference of the forces on the upper and lower faces, that is—

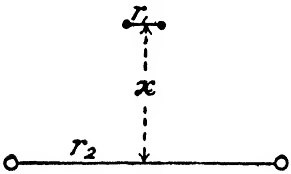


Fig. 234.

$$\frac{d\mathrm{H}}{dx}\,.\,dx\,.\,m\,.\,\pi r_1^2.$$

But $m\,.\,dx$ is the magnetic moment of unit area of the shell; that is, the strength σ of the shell.

$$\therefore\ \text{force}=\pi r_1^2\,.\,\sigma\,.\,\frac{d\mathrm{H}}{dx}.$$

But $\sigma=i_1$, the current in r_1.

$$\therefore\ \text{force}=\pi r_1^2 i_1\,.\,\frac{d\mathrm{H}}{dx}.$$

Now on p. 228 we showed that—

$$\mathrm{H}=\frac{2\pi r_2^2 i_2}{(r_2^2+x^2)^{\frac{3}{2}}},$$

$$\therefore\ \frac{d\mathrm{H}}{dx}=-\frac{6\pi r_2^2 i_2 x}{(r_2^2+x^2)^{\frac{5}{2}}},$$

$$\therefore\ \text{force}=\frac{6\pi^2 r_1^2 r_2^2 i_1 i_2\,.\,x}{(r_2^2+x^2)^{\frac{5}{2}}}.$$

This is evidently zero when $x=0$, and by differentiating it again we may show that it is a maximum when $x=\frac{r_2}{2}$.

Kelvin's Ampere Balance.—Although the forces between current circuits cannot in general be calculated by simple means, it follows from the equivalence of the circuits with magnetic shells, that the forces between them are always proportional to the product of the current strengths.

In the case of Kelvin's ampere balance, the forces between parallel circular coils are balanced against a gravitational force. The value of the current cannot be determined in absolute measure from the force and the dimensions of the coils, so that it is necessary to calibrate the instrument by means of a silver

voltameter. The four coils, A, B, C, D (Fig. 235), are fixed, and the two, E and F, are attached to the movable arm which also carries a horizontal scale on which the weight W slides. The coils are all connected in series in such a way that when the current flows, the forces between A and E, D and E, urge E downwards; similarly F is urged upwards. The movable arm is suspended by the conducting wires which bring the current to E and F, and the centre of gravity of the arm can be adjusted by means of a metal flag until, when there is no current, the arm is horizontal when the sliding weight W is on the zero mark. The couple due to the current can then be balanced by sliding W to the right along the arm, the couple being proportional to the displacement of the weight. Since the force between any pair of coils is proportional to the current in each, the down-

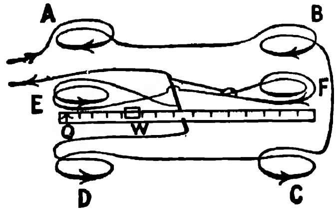


FIG. 235.

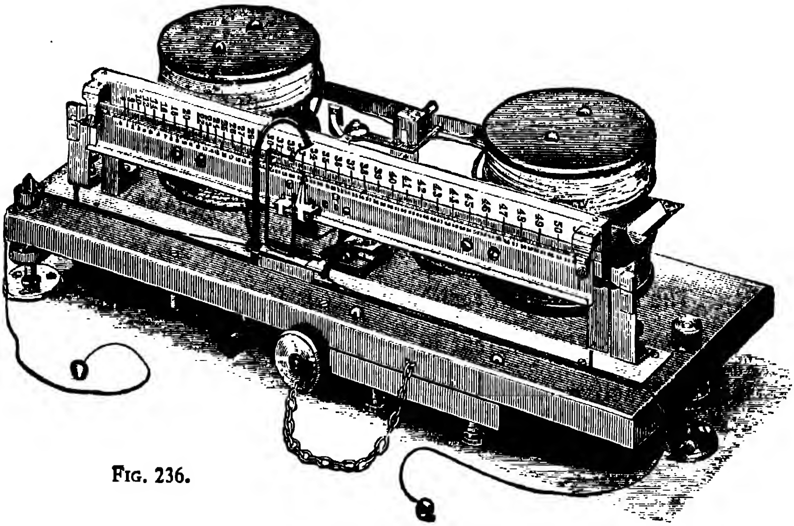


FIG. 236.

ward force on E and the upward force on F are each proportional to the square of the current in the instrument, and the couple is therefore proportional to i^2.

Thus if d is the displacement of W required to restore equilibrium on passing the current—

$$i^2 \propto d$$

$$\text{or, } i = k\sqrt{d}.$$

The constant k is determined when the instrument is cali-

brated, and a fixed scale is also attached which is marked directly in amperes.

There are several weights supplied with the instrument to alter the range, and for each weight a corresponding counterpoise also supplied must be placed in the tray at the end of the beam.

The general appearance of the instrument is shown in Fig. 236.

Kelvin Watt-Balance.—The watt-balance is similar in design to the ampere balance, but the movable coils E and F (Fig. 237) have high resistance and are not connected in series with the fixed coils. If the power absorbed in say a lamp L is required, the current in the lamp is caused to flow through the fixed coils A, B, C and D in series. The movable coils E and F are connected through a high resistance R_1 (to make the resistance up to, in some cases, 1000 ohms) to the points MN between which the power is being absorbed. Then if the current in the lamp is I amperes, this is also the current in the fixed coils, and if the

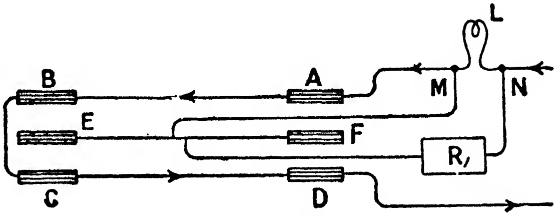


FIG. 237.

difference of potential between M and N is E volts, the current in the movable coils is $\frac{E}{R}$, R being their resistance together with R_1. The force between each pair of coils B–E, E–C, etc., being proportional to the current in each, is proportional to $\frac{IE}{R}$, and the couple acting on the beam, due to these forces being balanced as before by the displacement d of the weight, we have—

$$\frac{IE}{R} \propto d,$$

or, since R is constant,

$$IE = kd.$$

But IE is the power in watts absorbed in the lamp, and this is consequently proportional to the displacement of the movable weight required to maintain equilibrium. The constant k is determined by finding the displacement d for a known power, as measured by a standard ammeter and voltmeter, and the scale is usually graduated directly in watts. The adjustments are carried out as in the case of the current balance, and several

weights are supplied to enable the range of the instrument to be varied.

Siemens' Electro-Dynamometer.—Two coils, ABCD and *abcd*, are situated at right angles to each other, and when the instrument is used as an ammeter the coils are connected in series. With the connections as shown in Fig. 238, there is an attraction between AB and *ab* and also between CD and *cd*, the currents being in the same direction; but between AB and *cd*, and likewise between *ab* and CD, there are repulsions, and it will be noticed that all these forces tend to rotate the coil ABCD in the direction marked by the arrows, and further, that each of these forces is proportional to i^2. ABCD is suspended by a fibre and the light spiral spring S, which is attached to a pointer at the torsion head, and exerts a controlling couple, proportional to the twist in the spring. A pointer P is attached to the movable coil and serves as an indicator. This is in its equilibrium position for zero current. On passing the current, the coil is deflected, but is brought back to its zero position by rotating the torsion head, the amount of twist necessary to be put into the spring to effect this being measured by means of the circular scale.

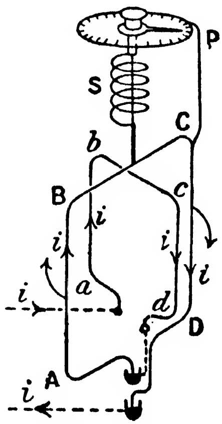


FIG. 238.

Then,
$$\begin{aligned} \text{couple} &\propto \text{twist}(=\theta) \\ &\propto i^2 \\ \therefore\ i^2 &\propto \theta, \text{ or, } i = k\sqrt{\theta}. \end{aligned}$$

The constant k may be found by observing θ for a known current, and the instrument may afterwards be used as an ammeter.

This instrument is sometimes designed for use as a wattmeter; the fixed coil having a great many turns of fine wire to ensure a high resistance. The low resistance coil is then placed in series with the circuit, the power absorbed in which it is required to measure, and the high resistance coil is placed in parallel across it. With this arrangement, current in series coil is I, and current in shunt coil $\frac{E}{R}$, as in the case of the Kelvin wattmeter (p. 248).

$$\therefore \text{couple} \propto \frac{EI}{R} \propto \theta,$$

$$\therefore EI = k\theta.$$

Thus the power absorbed in the circuit is directly proportional

to the twist in the spring necessary to maintain the movable coil in equilibrium at its zero position.

Electromagnetic Induction.—While attempting to find out whether a steady current produces another in neighbouring circuits, in a manner analogous to that in which electric charges are produced by the influence of other charges (p. 109), Faraday found that so long as the current is steady the result is negative, but on starting the current, a transient current in the opposite direction flows in the neighbouring circuit. The arrows in Fig. 239 indicate the directions of the transient currents in B when that in A is started and stopped. Exactly similar effects might be produced in B by advancing towards it from the side A, a bar magnet with its S pole facing B. The transient current in B is in the direction of that produced on starting the current in A. Similarly on withdrawing the magnet the effect is the same as that of stopping the current in A.

Faraday explained these results by stating that when the total

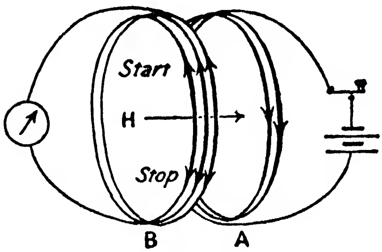


FIG. 239.

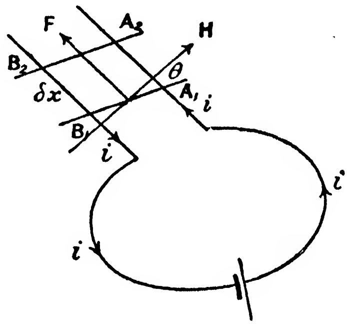


FIG. 240.

magnetic induction linked with a circuit changes, an electromotive force acts round the circuit, the direction of the electromotive force depending on the sign of the change of magnetic induction.

The actual value of the electromotive force due to a change in the magnetic flux linked with any circuit, may be deduced from our knowledge of the force acting on a circuit carrying current in a magnetic field, by making use of the principle of the conservation of energy. Consider a piece A_1B_1 of a circuit in which current i is flowing (Fig. 240), and let H be the magnetic field, making an angle θ with A_1B_1, and with the plane of the rails. Then the force per unit length of A_1B_1 is $Hi \sin \theta$ and is in the direction F at right angles to H and A_1B_1. Let A_1B_1 slide upon parallel conducting rails in the direction of this force. If length of A_1B_1 is l, work done for displacement δx, is—

$$F\delta x = Hil \sin \theta \, . \, \delta x.$$

Now, if e be the electromotive force of the battery maintaining the current i, work done in time δt is $ei\delta t$, and this is partly used in overcoming the resistance r of the circuit, the remainder being employed in moving the conductor A_1B_1. Now, work done in overcoming resistance is $i^2r\delta t$, and if there is no other action than these two in the circuit, we have by the principle of the conservation of energy—

$$ei\delta t = i^2r\delta t + \mathrm{H}il \sin\theta \,.\, \delta x,$$

$$\text{or, } i = \frac{e - \dfrac{\mathrm{H}l \sin\theta \,.\, \delta x}{\delta t}}{r}.$$

Thus the electromotive force e of the circuit is opposed by an electromotive force $\dfrac{\mathrm{H}l \sin\theta \,.\, \delta x}{\delta t}$.

On referring again to Fig. 240, we see that $l\delta x$ is the area described by the conductor in moving a distance δx, and $\mathrm{H} \sin\theta$ is the component of H perpendicular to this. Hence the product $(\mathrm{H} \sin\theta)(l\delta x)$ is the total normal magnetic flux over the area $A_1B_1B_2A_2$ when the medium is air; when the space has permeability μ, we must multiply by this amount, and in the above reasoning H must be replaced by B.

In any case, calling N the total normal flux over the whole circuit, $\mathrm{B}l \sin\theta \,.\, \delta x$ is the change in this amount ($\delta \mathrm{N}$) on account of the motion of A_1B_1, and we therefore see that this motion produces an electromotive force $\dfrac{\delta \mathrm{N}}{\delta t}$ in the circuit. In the limit when δt is infinitesimal—

$$\text{E.M.F. due to change of flux} = -\frac{d\mathrm{N}}{dt}.$$

The negative sign is taken because the electromotive force always opposes that producing the current, when the motion of the circuit is in the direction due to the electromagnetic actions themselves. If by some external agency the conductor were forced from A_2B_2 to A_1B_1 the direction of the induced electromotive force would be the same as e, but N is now diminishing, so that $\dfrac{d\mathrm{N}}{dt}$ is again negative.

Rule I.—The direction of the induced electromotive force is related to that of the motion and the magnetic field, in a manner illustrated by the three vectors in Fig. 241, the positions of which may be remembered by means of Prof. Fleming's *Right Hand Rule.* Extend the thumb, fore-finger and middle finger of the *Right* hand until they are mutually at right angles. Then if the

*F*ore-finger points along the magnetic *F*ield, the m*I*ddle finger along the current, I, the thu*M*b will then point in the direction of *M*otion.

Or, as an alternative, look along the direction of the magnetic field, then an anti-clockwise rotation brings the direction of the motion into that of the induced E.M.F. or current; that is, it brings the direction of the "cause" into that of the "effect." This rule has the advantage that it is identical with that for finding the direction of the force on a current in a magnetic field (see p. 240).

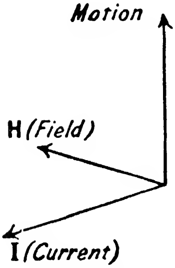


Fig. 241.

Rule II.—Another simple and useful rule for remembering the direction of the current in the whole circuit may be obtained by an inspection of Fig. 239. If the observer look along the magnetic lines of force towards the circuit, the induced current is anti-clockwise when the induction is increasing, and clockwise when it is diminishing.

If the circuit in Fig. 240 were considered to be flexible so that each element were movable, each part would travel outwards, the limit of travel being reached when the conductor became circular, in which case it would embrace the maximum amount of flux, and hence the rule given by Maxwell, that *a circuit always tends to move in that direction which tends to make the amount of magnetic flux through it a maximum.* This rule is sometimes of great convenience in determining the direction of a force acting on a circuit due to a magnetic field.

It should be noted that if the direction of H in Fig. 240 be reversed, the direction of motion is reversed. The circuit, if flexible, will then shrink and will eventually turn over and expand in the opposite direction, the motion all the time being in a direction towards the condition for the embracing of maximum magnetic flux by the circuit. This may easily be shown by taking a loop of thin, flexible, rubber-covered wire and tying a piece of thread round at a distance of about 10 cm. from the end of the loop. On hanging it between the poles of an electromagnet (Fig. 242) the loop will spread out to an approximately circular form on passing a current of a few amperes round it. If the current be suddenly reversed, the loop collapses, and expands in the opposite direction, always reaching equilibrium when the current is clockwise, as seen from the N pole of the magnet.

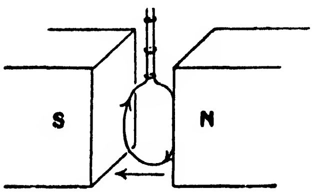


Fig. 242.

Lenz's Law.—Another generalisation on the laws of electromagnetic induction is due to Lenz,[1] which states that when a conductor moves with respect to a magnetic field, the currents induced in the conductor are in such a direction that the reaction between them and the magnetic field opposes the motion.

This law follows at once from the principle of the conservation of energy ; for if the forces due to the motion were in any other direction the motion would be increased, and it would only be necessary to start a conductor moving in a magnetic field and its velocity would continually increase, which is contrary to experience.

We can easily see that Lenz's law follows from the electromagnetic effects that we have already studied. For if the magnet and the conducting loop (Fig. 243) approach each other, the induced current in the loop as seen from the magnet is anti-clockwise (Rule II, p. 252), since the induction is increasing. Hence the equivalent magnetic shell has its N polar side towards the magnet, and there is consequently a repulsion between them. Their relative motion is thus opposed. If the direction of motion is reversed, the effects are all reversed and an attraction results, which is again in accordance with Lenz's law.

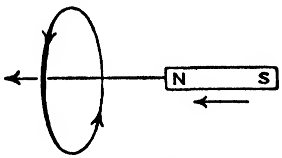


FIG. 243.

A well-known experiment in which a copper disc is caused to rotate underneath a suspended magnet, the magnet then being dragged round in the direction of rotation of the disc, is easily explained by the electromagnetic effects. For the motion of the conductor in the magnet's field produces currents which tend to prevent the relative motion of the magnet and the disc, and the magnet therefore follows the disc. This is known as Arago's disc experiment.

If the disc were delicately suspended and the magnet caused to rotate, a similar explanation would show that the disc would follow the magnet. This is the principle upon which the polyphase induction motor is founded, a rotating magnetic field produced by alternating currents, causing a closed conductor mounted upon an axle to rotate in the direction of rotation of the field (Chap. XI).

Use is made of Lenz's principle in constructing galvanometers of a dead-beat type, in which the suspended needle or coil will quickly come to rest after being disturbed. With an undamped system, the oscillations that occur after every movement render it exceedingly tedious to measure deflections, or to find the zero position in making Wheatstone's bridge tests. The oscillations

[1] Lenz, *Ann. de Phys.*, 81, p. 483. 1834.

are therefore damped out by surrounding the needle by a copper enclosure, the reaction between the induced currents in the enclosure and the needle itself tending to destroy the motion of the needle.

In the case of the suspended coil instrument a copper ring is placed inside the coil, the induced currents in which, as the coil oscillates, quickly bring it to rest. The presence of the ring does not in any way disturb the position of equilibrium when making readings of deflections, unless there are magnetic impurities in it, as the currents only exist when the ring is moving.

When unprovided with a damping ring, the suspended coil may be quickly brought to rest in its zero position by short-circuiting the galvanometer terminals, the induced currents taking place in the coil itself.

Circulation of Charge due to Induced Electromotive Force.—In a closed circuit, the electromotive force produced by the variation of the magnetic flux linked with the circuit we have seen to be $-\frac{dN}{dt}$. The current due to this being i, we have $i=-\frac{1}{r}\cdot\frac{dN}{dt}$, where r is the resistance of the circuit. In the interval of time dt, the amount of charge crossing any section of the circuit is $idt=dq$, since

$$i=\frac{dq}{dt} \qquad \text{(see p. 120)}.$$

$$\therefore\ dq=-\frac{1}{r}\cdot\frac{dN}{dt}dt,$$

$$=-\frac{dN}{r}.$$

Therefore the whole charge q passing any section as the magnetic flux linked with the circuit changes from zero to N, is—

$$q=-\int_0^N\frac{dN}{r}=-\frac{N}{r}.$$

If the flux N is removed from the circuit, the total charge passing any section is $+\frac{N}{r}$. The sign indicates the direction in which the charge passes round the circuit. Its magnitude is in each case $\frac{N}{r}$.

Ballistic Galvanometer—Suspended Magnet Type.—In order to measure the quantity of charge q which passes each point in a closed circuit, a galvanometer is inserted in the circuit. But the instrument must be of a particular type, as it has to integrate the quantity idt, in order to give the charge q. If the suspended

needle or coil be free to execute simple harmonic oscillations, then an impulse given to it when at rest in its equilibrium position will set it oscillating, and the amplitude of the oscillation may be taken as a measure of the impulse.

The following simple theory determines the relation between the quantity of charge passing through the coil of the galvanometer and the resulting amplitude of oscillation. The method of treatment varies according to the type of instrument. We will take the suspended magnet form first. Let the current in the coil (Fig. 244) at any instant be i, then the magnetic field at the centre of the coil is Gi, where G is the field at the centre for unit current. The strength of pole of the suspended magnet being m, the force on it is Gim. This, acting for infinitesimal time dt, gives it an impulse $Gimdt$, and from start to finish of the current the total impulse is—

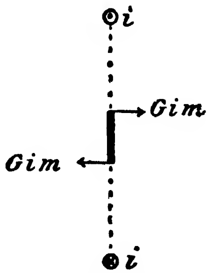


FIG. 244.

$$\int_0^T Gimdt = Gm\int_0^T idt.$$

But $\int_0^T idt = q$, the total charge sent round the coil ;

$$\therefore \text{ impulse} = Gmq.$$

If half the length of the magnet be l, moment of impulse about the centre is $Gmql$, and for both poles, $2Gmql$.

But $2ml = M$, the magnetic moment of the magnet ;

$$\therefore \text{ moment of impulse} = MGq.$$

Consequently this is the angular momentum of the magnet, which, if I be its moment of inertia and ω its angular velocity, is $I\omega$;

$$\therefore GMq = I\omega \quad . \quad . \quad . \quad . \quad . \quad . \quad (i)$$

If we could observe ω, then knowing the constants, G, M and I, we could calculate q. We cannot, however, observe ω directly, but it may be determined indirectly as follows.

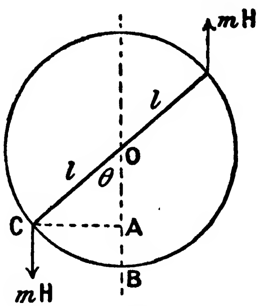


FIG. 245.

The impulse being over before the needle has rotated appreciably from its position of equilibrium, the needle possesses kinetic energy $\frac{1}{2}I\omega^2$ while still in the zero position. The needle will rotate until the controlling field brings it to rest, that is, until the work done in opposition to the controlling force is equal to the original kinetic energy. In the enlarged view of the needle in Fig. 245,

mH is the controlling force on each pole, and if θ is the angle of deviation when the needle has just lost all its kinetic energy and is on the point of turning back, the work that has been done by the force mH on the pole is—

$$mH \times AB = mHl(1-\cos\theta),$$

and for both poles,
$$\text{work} = 2mHl(1-\cos\theta)$$
$$= MH(1-\cos\theta),$$
$$\therefore\ MH(1-\cos\theta) = \tfrac{1}{2}I\omega^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

Eliminating ω from (i) and (ii),

$$\omega^2 = \frac{G^2M^2q^2}{I^2} = \frac{2MH(1-\cos\theta)}{I},$$

$$\therefore\quad q^2 = \frac{I}{MH} \cdot \frac{4H^2}{G^2} \cdot \frac{(1-\cos\theta)}{2}$$
$$= \frac{I}{MH} \cdot \frac{4H^2}{G^2} \cdot \sin^2\frac{\theta}{2}.$$

Now the time of oscillation T of the suspended needle, vibrating in a magnetic field, has been shown on p. 24 to be given by—

$$T = 2\pi\sqrt{\frac{I}{MH}}, \qquad \therefore\ \frac{I}{MH} = \frac{T^2}{4\pi^2}.$$

Hence,
$$q^2 = \frac{H^2T^2}{\pi^2G^2}\sin^2\tfrac{1}{2}\theta,$$
$$q = \frac{HT}{\pi G}\sin\tfrac{1}{2}\theta.$$

Ballistic Galvanometer—Suspended Coil Type.—In the case of the suspended coil galvanometer, the force on each vertical side of the coil for current i flowing in it is iHl, where H is the magnetic field due to the permanent magnet (see Fig. 226) and l the length of the vertical side of the coil. The impulse as before is $\int_0^T iHl\,dt = Hlq$, and the moment of momentum about the axis of suspension is $Hlbq$, for the two sides, where b is the length of a horizontal side of the coil. For n turns, each of area lb, the total effective area $A = nlb$, and the momentum equation is—

$$HAq = I\omega \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

As in the previous case, the kinetic energy is $\frac{1}{2}I\omega^2$, but now the coil is brought to rest by performing work in twisting the suspension.

If c be the restoring couple for unit twist in the suspension fibre, $c\theta$ is the couple for twist θ, and the work done for an additional small twist $d\theta$ is $c\theta d\theta$.

$$\therefore\ \text{whole work done in twisting suspension} = \int_0^\theta c\theta\,d\theta = \tfrac{1}{2}c\theta^2$$

where θ is the deviation of the coil when its kinetic energy $\frac{1}{2}I\omega^2$ is just expended in twisting the suspension.

$$\therefore \frac{1}{2}I\omega^2=\frac{1}{2}c\theta^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

$$\omega^2=\frac{c}{I}\theta^2.$$

From equation (i), $$\omega^2=\frac{H^2A^2q^2}{I^2},$$

$$\therefore \frac{H^2A^2q^2}{I^2}=\frac{c}{I}\theta^2,$$

$$q^2=\frac{c^2}{H^2A^2}\cdot\frac{I}{c}\theta^2.$$

Now the time of one torsional oscillation of a body of moment of inertia I when c is the restoring couple for unit angle of torsion is—

$$T=2\pi\sqrt{\frac{I}{c}}.$$

$$\therefore q^2=\frac{c^2T^2}{4\pi^2H^2A^2}\theta^2,$$

and, $$q=\frac{cT}{2\pi HA}\theta.$$

It will thus be seen that the relation between q and θ is not the same for the two types of galvanometer, the charge passing through the instrument being proportional to the sine of the angle of throw θ in the suspended magnet type, and to the angle θ itself in the suspended coil galvanometer. It should be noted that H in the first case is the controlling field; in the second it is the field to which the deflection is due, and hence it appears in the numerator in the first case and the denominator in the second. These two magnetic fields should not be confused with each other.

Damping.—It never happens that the vibration of the suspended system is simple harmonic; the vibrations always die away, the amplitude getting less and less. This decrease is due to a number of causes, the most important of which are the resistance of the air to the motion of the system, and the electromagnetic damping described on p. 254. On observing successive values of θ to left and right as the needle swings, it will be found that the ratio of one value to the next is a constant. Taking then θ_1, θ_2, θ_3 etc. as the succeeding values of θ, we have—

$$\frac{\theta_1}{\theta_2}=\frac{\theta_2}{\theta_3}=\frac{\theta_3}{\theta_4}= \quad . \quad . \quad . \quad =d.$$

This constant ratio d is called the decrement, and $\log_\epsilon d$ the

logarithmic decrement λ; therefore $\log_\epsilon d=\lambda$ and $d=\epsilon^\lambda$. When the oscillation is simple harmonic it may be represented by the projection of a rotating vector OA (Fig. 246) upon any fixed straight line, say OB. In this case succeeding amplitudes are OB, OB′, OB etc.; but if the amplitudes get less in a constant ratio, we must imagine the rotating vector OA to shrink at a constant rate, and it will be seen from the diagram that the shrinkage from θ_1 to θ_2 takes place in half a vibration. If there were no damping, the amplitude would all the time have been $\theta=$OA; and since the impulse was given to the system when in its middle position, that is the position corresponding to OA, the shrinkage of the vector that has actually occurred, before the first throw θ_1 is observed, has taken place during a quarter of a vibration.

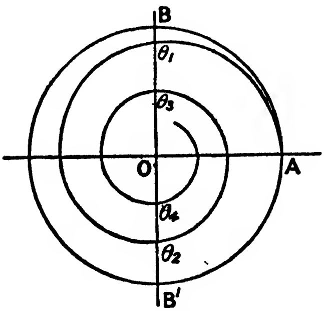


FIG. 246.

Now, for half a vibration, shrinkage is—

$$\frac{\theta_1}{\theta_2}=\frac{\theta_2}{\theta_3}=\frac{\theta_3}{\theta_4}= \ . \ . \ . \ =\epsilon^\lambda,$$

and for a whole vibration $\frac{\theta_1}{\theta_3}=\epsilon^{2\lambda}$, and so on, the shrinkage being proportional to the power of ϵ; and hence for a quarter vibration it is $\epsilon^{\frac{\lambda}{2}}$. Therefore $\frac{\theta}{\theta_1}=\epsilon^{\frac{\lambda}{2}}$,

$$\theta=\theta_1\epsilon^{\frac{\lambda}{2}}=\theta_1\left(1+\frac{\lambda}{2}+\frac{\lambda^2}{4\underline{|2}}+\frac{\lambda^3}{8\underline{|3}}+ \ . \ . \ .\right).$$

Since d must always be nearly equal to unity for a ballistic galvanometer, λ is always very small, and λ^2 and the higher powers of λ may be neglected.

$$\therefore \ \theta=\theta_1\left(1+\frac{\lambda}{2}\right).$$

Thus we can correct for the damping of the needle, although we cannot avoid it, and the equation for the two types of ballistic galvanometer will then be—

$$q=\frac{\text{HT}}{\pi\text{G}}\sin\tfrac{1}{2}\theta\left(1+\frac{\lambda}{2}\right),$$

and,

$$q=\frac{c\text{T}}{2\pi\text{HA}}\theta\left(1+\frac{\lambda}{2}\right).$$

If great accuracy is not required, the undamped throw may be obtained from the damped throw by multiplying it by $\sqrt{d}$, where d is the decrement, obtained by observing two consecutive elongations.

$$\text{For, } \frac{\theta}{\theta_1}=\epsilon^{\frac{\lambda}{2}}, \qquad \therefore \left(\frac{\theta}{\theta_1}\right)^2=\epsilon^{\lambda}=d,$$

$$\text{and, } \theta=\theta_1\sqrt{d}.$$

The advantage in the longer method lies in the fact that a great number of swings may be taken to determine λ, for if θ_1 and θ_{11} are observed, 10 half-vibrations occur between the observations—

$$\frac{\theta_1}{\theta_{11}}=\epsilon^{10\lambda}, \qquad \therefore \lambda=\tfrac{1}{10}\log_\epsilon\frac{\theta_1}{\theta_{11}}.$$

If a galvanometer is very heavily damped, the maximum of the swing occurs at much less than a quarter period from the start. Such a galvanometer must not be used for ballistic work.

Calibration of Ballistic Galvanometer.—In using the ballistic galvanometer to compare charges, or magnetic fluxes, the ratio only of the two respective throws is required, and the constants occurring in the equations need not be found. If however the charge or the flux is required in absolute measure, we must by some means determine these constants. The most convenient method is to pass a steady current through the galvanometer by means of a standard cell, a high resistance being included in the circuit. If a sufficiently high resistance is not available, a known fraction of the electromotive force of the cell may be obtained by means of a resistance box used as a shunt S (Fig. 247). If then the effective electromotive force applied to the galvanometer circuit be e, and r the resistance of the circuit, current $i=\frac{e}{r}$.

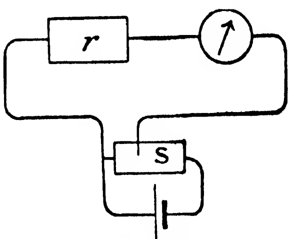


FIG. 247.

Then in the case of the suspended magnet galvanometer there will be a steady deflection θ_1, where $\frac{Gi}{H}=\tan\theta_1$, that is,

$\frac{Ge}{rH}=\tan\theta_1$,

$$\therefore \frac{G}{H}=\frac{r\tan\theta_1}{e},$$

and the equation for q on p. 256 becomes—

$$q=\frac{eT}{\pi r}\cdot\frac{\sin\frac{1}{2}\theta}{\tan\theta_1},$$

which may be written $q=\frac{eT}{2\pi r}\cdot\frac{\theta}{\theta_1}$, when θ and θ_1 are small.

For the suspended coil galvanometer, the steady deflection θ_1 is given by $i\mathrm{AH}=c\theta_1$(p. 241), that is $\frac{e\mathrm{AH}}{r}=c\theta_1$, and substituting $\frac{r\theta_1}{e}$ for $\frac{\mathrm{AH}}{c}$ in the equation for q on p. 257, we have $q=\frac{eT}{2\pi r}\cdot\frac{\theta}{\theta_1}$.

We see that when the calibration is performed in this way we are led to identical equations for the quantity of charge passing through the galvanometer with both types of instrument. T is determined by observing the time for a number of complete oscillations.

Capacities.—If the condenser C be charged by means of the cell of electromotive force e_1, the charge on the condenser will be $e_1\mathrm{C}=q$, and if this be sent through the ballistic galvanometer, we have from the last equation—

$$e_1\mathrm{C}=q=\frac{eT}{2\pi r}\cdot\frac{\theta\left(1+\frac{\lambda}{2}\right)}{\theta_1},\ \text{or},\ \mathrm{C}=\frac{eT}{2\pi r e_1}\cdot\frac{\theta\left(1+\frac{\lambda}{2}\right)}{\theta_1}.$$

If the electromotive force e_1 used for charging the condenser be the same as e, that used in calibrating the galvanometer, the expression for the capacity becomes $\mathrm{C}=\frac{T}{2\pi r}\cdot\frac{\theta}{\theta_1}$. It should be noted that if the resistance is given in ohms, C will be in farads (see Chap. XII).

The charge and discharge key K (Fig. 248) is a useful one; on depressing it the condenser is charged, and on releasing it, first one end of the battery is insulated, and then the galvanometer circuit is closed so that the condenser is discharged through it.

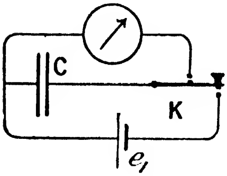


FIG. 248.

When a comparison of two capacities merely is wanted, it is not necessary to calibrate the galvanometer; we obtain a throw θ_1 by discharging the first condenser C_1 through the galvanometer, having previously charged it by means of a cell of electromotive force e_1 (Fig. 248), and then repeat the process for the second condenser C_2, obtaining the throw θ_2.

Then $C_1=k\theta_1$, and, $C_2=k\theta_2$,

$$\therefore \frac{C_1}{C_2}=\frac{\theta_1}{\theta_2},$$

since the other quantities in the multiplier $\left(k=\frac{HT}{\pi G}, \text{ or } k=\frac{cT}{2\pi AH}\right)$ are constant, and therefore k is constant.

If one of the capacities be a standard, and the cell have known electromotive force, the instrument may be calibrated by finding the constant k from the relation $q=eC=k\theta$, but it should be noticed that the constant determined in this way may only be used when the galvanometer is on open circuit. When the circuit is closed, the resistance is not the same, and in general the time of oscillation T will be altered.

Resistance by Method of Damping.—In the case of a ballistic galvanometer, the resistance to the motion of the moving part when the circuit is open, is due to air friction, viscosity in the fibre, and to the induced currents in any neighbouring masses of metal. This whole effect is very small in a well-designed instrument, and we may consider the effect to be a couple, whose value at any instant is proportional to the angular velocity, and may therefore be written $p\cdot\frac{d\theta}{dt}$, which opposes the motion of the suspended part. When, however, the circuit is closed, the motion of the suspended needle causes the coil to be cut by a magnetic flux, and a current to be induced in it proportional to the angular velocity, and inversely as the resistance of the circuit, and the reaction between this and the permanent field gives rise to a retarding couple which we may write $\frac{m}{R}\cdot\frac{d\theta}{dt}$, where m is a constant involving the magnetic flux due to the magnet and the area of the coil.

The equation of motion of the suspended part may then be written—

$$I\frac{d^2\theta}{dt^2}+\left(\frac{m}{R}+p\right)\frac{d\theta}{dt}+c\theta=0,$$

where I is the moment of inertia of the moving part, and c the controlling couple per unit deflection when this is small.

Thus, $$\frac{d^2\theta}{dt^2}+\frac{1}{I}\left(\frac{m}{R}+p\right)\frac{d\theta}{dt}+\frac{c}{I}\theta=0,$$

or, $$\frac{d^2\theta}{dt^2}+2b\frac{d\theta}{dt}+k^2\theta=0,$$

when, $\frac{1}{I}\left(\frac{m}{R}+p\right)=2b$, and, $\frac{c}{I}=k^2$.

The solution of which is (see p. 337)—

$$\theta=A\epsilon^{-bt}\cos\sqrt{k^2-b^2}t.$$

Thus the elongation θ_0 when $t=0$, is A; and θ_1 when $t=\frac{\pi}{\sqrt{k^2-b^2}}$, is $-A\epsilon^{-b\frac{\pi}{\sqrt{k^2-b^2}}}$, and so on. Successive elongations being in opposite directions, the opposite signs of alternate elongations will be omitted, and we have—

$$\frac{\theta_0}{\theta_1}=\frac{1}{\epsilon^{-b\frac{\pi}{\sqrt{k^2-b^2}}}},\qquad \frac{\theta_1}{\theta_2}=\frac{\epsilon^{-b\frac{\pi}{\sqrt{k^2-b^2}}}}{\epsilon^{-b\frac{2\pi}{\sqrt{k^2-b^2}}}}=\frac{1}{\epsilon^{-b\frac{\pi}{\sqrt{k^2-b^2}}}},$$

$$\therefore\ \frac{\theta_0}{\theta_1}=\frac{\theta_1}{\theta_2}=\frac{\theta_2}{\theta_3}=\epsilon^{\frac{b\pi}{\sqrt{k^2-b^2}}}=d.$$

Thus, $\log_\epsilon d=\lambda=\frac{b\pi}{\sqrt{k^2-b^2}}=$logarithmic decrement.

Now, for the galvanometer to be ballistic, b must be small in comparison with k, so that we have—

$$\lambda=\frac{b\pi}{k}=\pi\frac{1}{2I}\left(\frac{m}{R}+p\right)\sqrt{\frac{I}{c}}$$

$$=a\left(\frac{1}{R}+p_1\right),$$

where $p_1=\frac{p}{m}$, and a is a new constant.

Let the logarithmic decrement λ_0 be determined with the galvanometer on open circuit; then $R=\infty$,

and, $\lambda_0=ap_1.$

Again, let it be determined with the galvanometer short-circuited, R_g being the resistance of the galvanometer itself.

Then, $\lambda_g=a\left(\frac{1}{R_g}+p_1\right),$

and again, with a total resistance R in circuit—

$$\lambda_R=a\left(\frac{1}{R}+p_1\right).$$

Subtracting λ_R from λ_g we get—

$$\lambda_g-\lambda_R=a\left(\frac{1}{R_g}-\frac{1}{R}\right).$$

Subtracting λ_0 from λ_R we get—

$$\lambda_R-\lambda_0=\frac{a}{R}.$$

Now, dividing the last equation but one by the last—

$$\frac{\lambda_0-\lambda_R}{\lambda_R-\lambda_0}=\frac{R-R_0}{R_0}.$$

$R-R_0$ is the resistance added for the third determination, and hence it can be found in terms of R_0 or *vice versâ.*

This method must be modified in the case of a moving coil galvanometer, for on short-circuiting the coil, the motion becomes dead beat; but two high resistances may be compared; for if R_1 and R_2 are the total resistances, including that of the galvanometer, and λ_1 and λ_2 the corresponding logarithmic decrements—

$$\lambda_0=ap_1,$$

$$\lambda_1=a\left(\frac{1}{R_1}+p_1\right),$$

$$\lambda_2=a\left(\frac{1}{R_2}+p_1\right),$$

From which as before—

$$\frac{\lambda_1-\lambda_2}{\lambda_2-\lambda_0}=\frac{R_2}{R_1}-1,$$

$$\text{or, } \frac{R_2}{R_1}=\frac{\lambda_1-\lambda_0}{\lambda_2-\lambda_0}.$$

Measurement of Magnetic Flux.—If a closed coil is rotated in a magnetic field, current flows in it owing to the electromotive force produced by the change in the magnetic flux linked with the coil. Let the coil have effective area A and the magnetic field in which it is situated be uniform and of strength H. Then, if the plane of the coil make angle θ with the direction of the field (Fig. 249)—

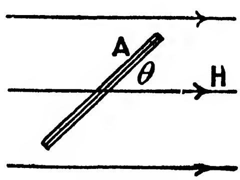


FIG. 249.

$$\text{magnetic flux linked with coil}=\text{HA}\sin\theta=\text{N}.$$

Hence, as the coil rotates—

$$e=-\frac{d\text{N}}{dt}=-\frac{d(\text{HA}\sin\theta)}{dt},$$

and at each instant the current i is $\frac{e}{r}$.

$$\therefore i=-\frac{\text{HA}}{r}\cdot\frac{d(\sin\theta)}{dt},$$

or, $$idt=-\frac{\text{HA}}{r}d(\sin\theta).$$

If, then, $\theta=\frac{\pi}{2}$ at time $t=0$; and $-\frac{\pi}{2}$ at the end of an interval of time t—

$$\int_0^t idt=q=-\frac{\text{HA}}{r}\int_{\theta=\frac{\pi}{2}}^{\theta=-\frac{\pi}{2}}d(\sin\theta),$$

$$=-\frac{\text{HA}}{r}\Big[\sin\theta\Big]_{\frac{\pi}{2}}^{-\frac{\pi}{2}}=\frac{2\text{HA}}{r}.$$

If, then, the coil is in series with the ballistic galvanometer, calibrated in the usual way—

$$\frac{2\text{HA}}{r}=k\theta,$$

or,

$$\text{H}=\frac{rk}{2\text{A}}\theta,$$

where k is the constant determined by the ordinary method (p. 259).

When H is the horizontal component of the earth's magnetic field, the coil is known as the *Earth Inductor*, Fig. 250, and the method may be used for determining H. On the other hand, if H is accurately known, the method is a convenient one for calibrating the ballistic galvanometer.

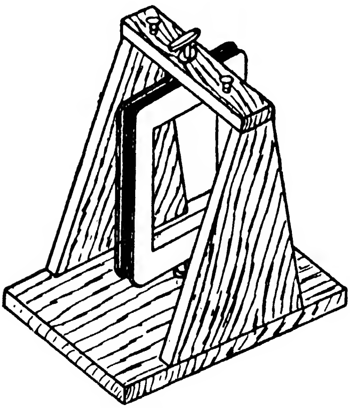


Fig. 250.

Care must be taken that before and after rotation, the coil shall be at right angles to the meridian, as otherwise the charge passing round the coil is less than $\frac{2\text{HA}}{r}$. The correct position is that in which a maximum throw is obtained for a sudden rotation of the coil through 180°.

The vertical component of the earth's field may be found by laying the apparatus on its side so that the coil is horizontal before and after rotation. Then $\text{V}=\frac{rk}{2\text{A}}\theta_{\text{V}}$, and the dip may be found by taking the ratio of the throws produced in the two positions—

$$\text{H}=\frac{rk}{2\text{A}}\theta_{\text{H}},\qquad \text{V}=\frac{rk}{2\text{A}}\theta_{\text{V}},$$

$$\tan(\text{dip})=\frac{\text{V}}{\text{H}}=\frac{\theta_{\text{V}}}{\theta_{\text{H}}}.$$

It should be noticed that in measuring a magnetic flux, the total amount of charge caused to circulate in the circuit is $\frac{N}{r}$, where r is the resistance of the circuit, and consequently to get the greatest effect on the galvanometer, r should be as small as possible. With the suspended coil galvanometer, r must not be indefinitely reduced, for this would have the effect of rendering the galvanometer dead beat, whereas the relation between charge and throw depends on it being ballistic. Hence for the measurement of small values of the flux it is desirable to use a suspended magnet galvanometer, which, even when short-circuited, is usually sufficiently ballistic for the purpose. For large values of the flux, a suspended coil galvanometer may be used, a high resistance being put into the circuit without unduly diminishing the sensitiveness; in fact, for measurements of magnetic permeability (p. 276) it is generally necessary to reduce the sensitiveness in this way, the additional advantage of rendering the galvanometer ballistic being attained.

In the measurement of capacity no such difficulty arises, for the amount of charge caused to pass through the galvanometer is independent of the resistance of the circuit. Hence the greater the number of turns in the coil of the galvanometer the better, and the sensitiveness of a high-resistance galvanometer is therefore greater than that of one of low resistance.

Standards of Magnetic Flux.—Standard magnetic fluxes are very useful for the calibration of ballistic galvanometers. One form of such standard consists of a long solenoid, on the middle part of which is wound a secondary coil of a great many turns.

If the solenoid has n_1 turns per centimetre length, the magnetic field in the interior when a current i absolute units flows in it is $4\pi n_1 i$ (p. 229). If then A be the mean area of section of the solenoid, the total magnetic flux across any section far removed from the end is $4\pi n_1 i A$. If the secondary coil has n_2 turns the effective amount of magnetic flux linked with it is $4\pi n_1 i A n_2$. This flux enters it on establishing the current in the solenoid, and leaves it on stopping the current, and if the ballistic galvanometer be connected to the secondary terminal S_1S_2 (Fig. 251), the charge caused to circulate through the galvanometer on starting or stopping the current is $\frac{4\pi n_1 n_2 A i}{r}$ absolute units, where r is the resistance of the circuit. Since all these quantities are easily measurable, the method is a convenient one for calibrating the galvanometer.

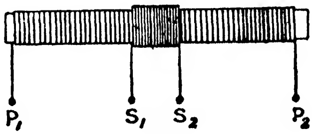


FIG. 251.

Another convenient source of flux is the Hibbert magnetic standard (Fig. 252). A block of hard steel has a cylindrical slot cut in it. The steel is magnetised so that the magnetic flux cuts radially across the slot. A circular coil C wound on a hollow brass cylinder can be dropped into the slot, and in doing so cuts the magnetic flux due to the cylindrical magnet. The number of magnetic lines must in the first instance be determined by comparison with some such standard as that last described, and it is found that owing to the form of the permanent magnet the flux remains very constant and forms a useful and easily employed standard.

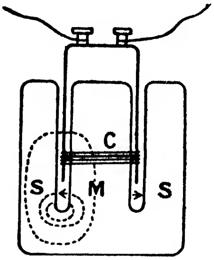


FIG. 252.

Grassot Fluxmeter.—The passage of the electric charge through the ballistic galvanometer must have ceased before the moving system has moved appreciably, or the impulse will not be applied to the system in its position of equilibrium, and our equations no longer apply. For measuring magnetic fluxes, the Grassot fluxmeter has the great advantage over the ballistic galvanometer, that the change in flux need not take place instantaneously, or at any particular rate, the moving coil being at rest before and after the change, the difference in position being proportional to the change in flux linked with the circuit. This result is attained by reducing to a very small amount all sources of damping other than that due to the electromagnetic effect between the permanent field and the coil as it rotates, until this becomes the predominating control. The coil BB (Fig. 253) is suspended by a single silk fibre attached to the spiral spring R to prevent damage from shocks. The current enters and leaves the coil by two fine silver spirals, S and S′. The mechanical control is thus very small, and the damping due to the air resistance to motion is usually insignificant, so that the only effective damping is that due to the induced current in the coil as it rotates, in fact the period of oscillation of the coil on open circuit is of the order of a minute.

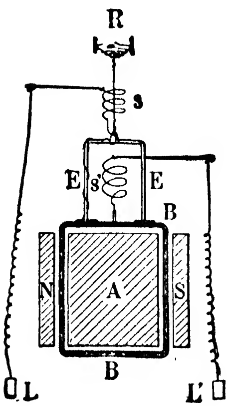


FIG. 253.

The terminals L and L′ are connected to an exploring coil, the variation in flux through which it is required to determine. If the effective area of the exploring coil be known, the magnetic induction can be calculated from the flux ($N = BA$). With the exploring coil connected, the total

resistance is of the order of 20 ohms, and the coil will remain practically at rest in any position.

When the flux through the exploring coil changes, there will be an electromotive force acting in the circuit, and consequently a current in it, whose value is $\frac{e}{r}=i$, where r is the resistance of the circuit and e the resultant electromotive force. The coil therefore experiences a couple iAH, A being the effective area of the galvanometer coil and H the field due to the permanent magnet.

$$\therefore\ i\mathrm{AH}=\mathrm{I}\frac{d\omega}{dt}$$

where I is the moment of inertia of the moving coil and ω its angular velocity, so that $\frac{d\omega}{dt}$ is its angular acceleration.

Then,
$$\frac{e\mathrm{AH}}{r}=\mathrm{I}\frac{d\omega}{dt}\quad\ldots\ldots\ldots\quad\text{(i)}$$

Now e is the difference between the electromotive force due to rate of change of flux in the exploring coil $\left(\frac{d\mathrm{N}}{dt}\right)$ and that in the galvanometer coil due to its rotation with angular velocity ω in the field of the permanent magnet. The latter is AHω, or AH$\frac{d\theta}{dt}$, where θ is the angle the coil makes with its mean position.

Thus,
$$e=\left(\frac{d\mathrm{N}}{dt}-\mathrm{AH}\omega\right),$$

and from (i),
$$\frac{\mathrm{AH}}{r}\left(\frac{d\mathrm{N}}{dt}-\mathrm{AH}\frac{d\theta}{dt}\right)=\mathrm{I}\frac{d\omega}{dt}.$$

Integrating, we have—

$$\frac{\mathrm{AH}}{r}\int_0^t\left(\frac{d\mathrm{N}}{dt}-\mathrm{AH}\frac{d\theta}{dt}\right)dt=\mathrm{I}\int_0^t\frac{d\omega}{dt}dt.$$

Now the last integral is $\Big[\omega\Big]_{t=o}^{t=t}$, and since the coil is at rest before and after the change in flux, both these limiting values of ω are zero.

$$\therefore\ \int_0^t\left(\frac{d\mathrm{N}}{dt}-\mathrm{AH}\frac{d\theta}{dt}\right)dt=0,$$

or,
$$\Big[\mathrm{N}\Big]_0^t-\mathrm{AH}\Big[\theta\Big]_0^t=0.$$

i.e.
$$\mathrm{N}=\mathrm{AH}\theta=k\theta$$

where N is the total change in magnetic flux linked with the exploring coil, and θ is the corresponding change in position of the suspended coil. This relation is independent of the time during which the change in flux takes place. Thus to measure a flux, all that is necessary is to put the exploring coil into the space in which the flux is required, in performing which action the coil cuts the flux to be measured, and the displacement of the galvanometer coil measures at once this flux. The coil is provided with a pointer moving over a scale so calibrated that with an exploring coil of definite resistance the flux for one division of the scale is known. A mirror is also attached to the moving system, so that by means of the deflection of a spot of light, very small magnetic fluxes may be measured, but in this case the scale must be calibrated by using a known flux to determine the constant k in the expression $N=k\theta$.

CHAPTER IX

MAGNETIC PROPERTIES OF MATERIALS

Theories of Magnetisation.—Passing over the early theories of magnetisation, which accounted for the phenomena by the existence of two magnetic fluids (Coulomb), and others by means of vortices (Decartes), we come to the first approach to a molecular theory, due to Poisson, who supposed that the magnetic materials contained small spheres which are conductors of the magnetic fluids, and in a magnetic field behave in an analogous manner to that of conducting spheres in an electric field. The next advance was due to Weber, who assumed that the molecules of a magnetic substance are themselves permanent magnets, and that in the act of magnetisation they are turned into the direction of the magnetising field. In order to account for the fact that a field, however weak, will not set all the molecular magnets parallel to the field, and therefore produce saturation, a mechanical restraint opposing their rotation was postulated. Sir J. A. Ewing added to the molecular theory by showing that the magnetic interaction between the molecules themselves is sufficient to account for the known behaviour of magnetic materials. It is now firmly established that magnetic poles have no existence, apart from electric currents or the motion of electric charges. Whatever the motion of electricity to which the magnetic properties of material are due, there is little doubt that it occurs within the atom. This motion is that of the electrons, but the form of motion does not concern us here, except that each motion is permanent, or nearly so (see Chap. XVI), and gives a magnetic moment to each atom. Nevertheless, the idea of magnetic pole has played a very useful role in the history of magnetism, and is still a very useful device in studying magnetic phenomena.

Intensity of Magnetisation and Magnetic Susceptibility.—We define the intensity of magnetisation (I) of a magnetised material as the *ratio of the magnetic moment to the volume of any piece of it*, the piece being sufficiently small for us to consider its magnetisation uniform.

Thus, $$I=\frac{\text{magnetic moment}}{\text{volume}}.$$

If a rectangular piece of the material (Fig. 254) has length l parallel to the direction of magnetisation and cross-section a, then if σ is the amount of pole per unit area of each end, we have, when the magnetisation is uniform,—total pole at each end is $a\sigma$, and the magnetic moment is $la\sigma$. But the volume is la,

$$\therefore\ \mathrm{I}=\frac{la\sigma}{la}=\sigma.$$

Thus the intensity of magnetisation may also be defined as the amount of pole per unit area taken at right angles to the direction of magnetisation.

If the volume taken be situated in the interior of a magnetised body, σ is not free pole, in the sense that it produces an outside effect, for it is situated indefinitely close to an equal and opposite amount of pole upon the adjacent layer. It is only where the magnetic poles form the outside layer of the material that their effects are not balanced by that of opposite poles, and the

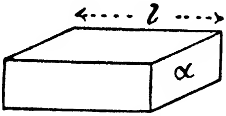


FIG. 254.

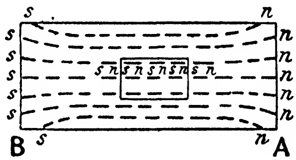


FIG. 255.

ordinary polar phenomena are produced, as at A and B (Fig. 255).

The ratio of the intensity of magnetisation (I) *at any point within a body to the magnetic field* (H) *to which it is due is called the magnetic susceptibility* (κ) *of the material.* Thus $\kappa=\frac{\mathrm{I}}{\mathrm{H}}$.

Magnetic Induction.—It will be seen that in dealing with magnetism the quantity I, or intensity of magnetisation, is analogous to the polarisation P in electrostatics (p. 115). Following the reasoning given on p. 116, it follows that in the interspaces between the atoms of a magnetic substance there is a magnetic intensity $\mathrm{H}+4\pi\mathrm{I}$ corresponding to $\mathrm{E}+4\pi\mathrm{P}$ in the electrostatic case (p. 116). As the quantity $\mathrm{E}+4\pi\mathrm{P}$ was called the electrical induction, ϕ, so $\mathrm{H}+4\pi\mathrm{I}$ will be called the *magnetic induction*, B,

that is,

$$\begin{aligned}\mathrm{B}&=\mathrm{H}+4\pi\mathrm{I}\\&=\mathrm{H}+4\pi\kappa\mathrm{H}\\&=\mathrm{H}(1+4\pi\kappa).\end{aligned}$$

By a substitution of κ for ϵ and I for P in the argument on p. 117, it may be shown that the force between two magnetic

poles m_1 and m_2 situated in a magnetic medium of susceptibility κ is $\frac{m_1 m_2}{(1+4\pi\kappa)r^2}$ the factor $(1+4\pi\kappa)$ involving 4π due to the effect of the magnetisation of the medium. As before, it is convenient to substitute a symbol for $1+4\pi\kappa$, and the one chosen is μ. It is called the *magnetic permeability* of the medium.

Thus, $$\mu=1+4\pi\kappa.$$

and, $$\text{force between poles}=\frac{m_1 m_2}{\mu r^2}.$$

It follows, from above, that $B=\mu H$.

The permeability μ is not, like the dielectric constant k, a constant for any one material. In the case of iron, nickel and cobalt the value of μ is much greater than for other substances and for any one specimen its value varies between wide limits.

Demagnetisation.—It must be understood that H in the above expression is the actual field producing magnetisation within the material, and that if there are free poles upon the specimen they will always produce a field which is in opposition to, and must be subtracted from, the original field, within the material, in order to obtain the resultant magnetising effect. Thus, for a magnet NS (Fig. 256) each pole produces its own radial field, the resultant being the ordinary field due to a pair of poles. At the middle of the magnet this field is opposed to the magnetising field H, and therefore exerts a demagnetising effect upon the bar. It is for the purpose of removing the free poles that produce this demagnetising effect, that permanent magnets are usually provided with soft-iron keepers, the keeper producing poles equal and opposite to those of the magnet, and being very nearly coincident in position with them, these poles produce a field equal and opposite to the demagnetising field.

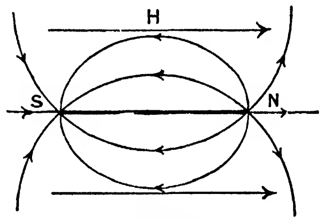


FIG. 256.

Whatever the form of the magnet, the demagnetising field is proportional to the strength of the pole to which it is due, and this in turn is proportional to the intensity of magnetisation, so that the demagnetising field is equal to NI, where N is a constant depending on the geometrical form of the magnetised body.

If then H′ is the magnetising field when the body is absent, and H that actually existing in the interior of the body—

$$H=H'-NI.$$

N may be calculated in a number of simple cases when the

interior field is uniform, but this is not in general the case. On p. 235 we saw that the resultant field inside a sphere of permeability μ_2 situated in a material of permeability μ_1 is $\frac{3\mu_1}{\mu_2+2\mu_1}$ times the external field. If now we consider *a sphere of iron of permeability* μ situated in air, the resultant interior field is $\frac{3}{2+\mu}$ times the original field, that is—

$$H=\frac{3}{2+\mu}H'.$$

Remembering that, $\mu=1+4\pi\kappa,$
and that, $I=\kappa H,$
we then have, $H=H'-\frac{4}{3}\pi I,$

from which we see that for a sphere, $N=\frac{4}{3}\pi$. Thus, with the value $\mu=1000$—

$$H=\tfrac{3}{1002}H',$$

and since $B=\mu H$—

$$B=\tfrac{3000}{1002}H'=3H' \text{ approximately.}$$

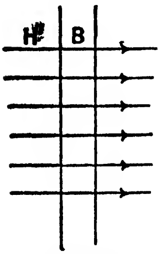


FIG. 257.

If there were no demagnetisation effect, B would have been 1000 H'; and hence the important part played by the free polar surfaces in this case.

The effect is greatest for a *magnetic sheet perpendicular to the field.* In this case the surface condition (ii), p. 234, tells us that $B=H'$ (Fig. 257).

Further, $B=\mu H,\qquad \therefore H=\frac{H'}{\mu}.$

But, $\mu=1+4\pi\kappa,\qquad \therefore H(1+4\pi\kappa)=H',$

$$H=H'-4\pi I.$$
$$\therefore N=4\pi.$$

For a permeability of 1000, the actual field inside the sheet is only $\frac{1}{1000}$ of that outside.

In the case of a *very long wire parallel to the field*, the demagnetisation effect throughout the greater part of the length of the wire is negligible, owing to the distance away of the poles at the ends. When the wire is not very long, Ewing [1] treated it as an ellipsoid with the axis parallel to the field of much greater length than the axes perpendicular to the field.

For an ellipsoid having semi-axes, a, b and c, when c is the long axis, parallel to the field, and $a=b=\sqrt{1-e^2}\,c$, it is shown in

[1] J. A. Ewing, "Magnetic Induction in Iron and other Metals."

"Maxwell's Treatise on Electricity and Magnetism," **vol. ii.** that—

$$N=4\pi\left(\frac{1}{e^2}-1\right)\left(\frac{1}{2e}\log\frac{1+e}{1-e}-1\right),$$

and using this equation, the values of N corresponding to different values of $\frac{c}{a}$, or ratio of length to diameter of the wire, are calculated—

$\frac{\text{length}}{\text{diameter}}$.	N.
50	0·01817
100	0·00540
200	0·00157
300	0·00075
400	0·00045
500	0·00030

Thus for a wire of length equal to 500 times its diameter—

$$H=H'-0{\cdot}00030I,$$

or,

$$\frac{H'}{H}=1+0{\cdot}00030\kappa.$$

κ does not often exceed 200, and for this value—

$$\frac{H'}{H}=1{\cdot}06.$$

Thus, the field is reduced about 6 per cent. by the free poles The magnitude of the effect shows that in measuring the susceptibility of an iron wire, it is usually necessary to correct the magnetising field for the demagnetising effect of the free poles upon the specimen.

Practical Methods.—(i) *Magnetometer.* For material in the form of a wire, the magnetometer, as developed by Ewing,[1] is usually employed in studying the magnetic properties. The specimen is placed vertically inside a magnetising solenoid, with its upper pole on a level with the needle of the magnetometer of the type shown in Fig. 7. If then a be the area of cross-section of the wire, the strength of pole at each end is Ia, when intensity of magnetisation is I, and the strength of field at the magnetometer needle M due to the pole A is $\frac{Ia}{d^2}$ (Fig. 258). That due to the pole B is $-\frac{Ia}{d^2+l^2}$, the horizontal component of which is

[1] J. A. Ewing, "Magnetic Induction in Iron and other Metals."

$-\frac{Ia\,.\,d}{(d^2+l^2)^{\frac{3}{2}}}$, and the resulting horizontal field due to the specimen is $Ia\left\{\frac{1}{d^2}-\frac{d}{(d^2+l^2)^{\frac{3}{2}}}\right\}$. When this is at right angles to the controlling field f, then the deflection θ, is given by—

$$Ia\left\{\frac{1}{d^2}-\frac{d}{(d^2+l^2)^{\frac{3}{2}}}\right\}=f\tan\theta.$$

If the specimen is very long the term $\frac{d}{(d^2+l^2)^{\frac{3}{2}}}$ is small, but in any case it may be calculated. The effective length l is about three-quarters of the length of the wire. If the controlling field f be known, I can then be found in terms of the deflection θ.

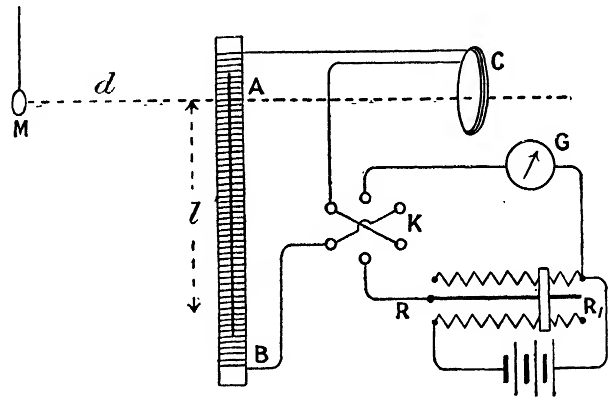


FIG. 258.

The magnetising field H′ is known in terms of the current i in the magnetising solenoid; n being the number of turns per centimetre length of coil, $H'=4\pi ni$, where i is in absolute units.

There are several disturbances to be allowed for. In the first place the solenoid produces a magnetic field at the magnetometer. This effect is eliminated by placing the vertical circular coil C, whose axis passes through A and M, in series with the solenoid and adjusting its distance from M, with the specimen removed, until on passing the current the magnetometer needle is undisturbed. This balance, if perfect for one current, holds for all currents, and the disturbing effect of the solenoid on the needle is then eliminated. The coil C serves another useful purpose, for if we disconnect the solenoid and observe the deflection θ_1 produced by a current i_1 in C, we can obtain the value of the controlling field f. Calling the distance of M from

C, x, the field at M due to the current in C is $\frac{2\pi n a^2 i_1}{(a^2+x^2)^{\frac{3}{2}}}$ (see p. 228), where n is the number of turns in C.

$$\therefore \frac{2\pi n a^2 i_1}{(a^2+x^2)^{\frac{3}{2}}} = f \tan \theta_1,$$

from which f can be found.

The second disturbance is due to the fact that the specimen is always magnetised by the vertical component of the earth's magnetic field. To eliminate this effect a second solenoid (not shown in the diagram) is wound upon the first and the current in it adjusted until, on demagnetising the specimen, the magnetometer needle remains in its true zero position when there is no magnetising current. This earth neutralising current is maintained constant during the experiment.

Having made all the adjustments, a series of values of θ and i is observed, beginning with the slider of the rheostat in the position R_1, so that the value of the magnetising current is small. The current is then increased step by step to a maximum by moving the contact from R_1 to R, θ and i being observed at each step. The current is then diminished in a similar manner to zero, then reversed by means of the key K, increased to a negative maximum, diminished to zero, and finally reversed and increased to its original positive maximum. The series of readings of θ and i are then converted by constant factors, determined as above described, into the corresponding values of I and H′, which may then be plotted in the form of a curve. In the case of the ordinary reflecting magnetometer, the deflection is usually sufficiently small to use θ instead of $\tan \theta$, without appreciable error.

Then—

$$\mathrm{I} = \frac{2\pi n a^2 i_1}{(a^2+x^2)^{\frac{3}{2}}\theta_1} \cdot \left\{\frac{1}{d^2} - \frac{d}{(d^2+l^2)^{\frac{3}{2}}}\right\}^{-1} \cdot \frac{1}{a} \cdot \theta.$$

$$\mathrm{H}' = 4\pi n i.$$

If the currents are measured in amperes, each of these expressions must be divided by 10.

The dotted curve H′ for a specimen of steel piano wire is shown in Fig. 259. To obtain the curve connecting I and H from this, the demagnetisation effect is to be allowed for. To do this the line Oq is drawn through the origin, making angle pOq such that the $\tan pOq = \mathrm{N}$. But $\mathrm{H} = \mathrm{H}' - \mathrm{NI}$. Then,

$$pq = Op \tan (qOp) = Op \cdot \mathrm{N} = \mathrm{NI}.$$

Thus, $pq = \mathrm{H}' - \mathrm{H}$, and drawing $ef = pq$, horizontally from the

point e of the H′ curve, we get the corresponding point f on the H curve. The whole curve is then corrected in the same way. This is equivalent to shearing the H′ curve through an angle $\tan^{-1}$ N to obtain the true H curve, and obviates the necessity of calculating the demagnetisation effect H′—H or NI for every reading.

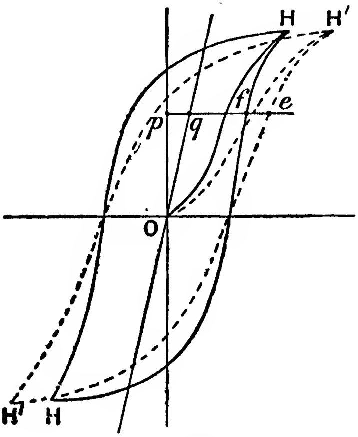


FIG. 259.

(ii) *Ballistic Method.*—In this method the total magnetic flux in the material is measured by means of the ballistic galvanometer. The material in the form of a ring, usually of circular cross-section, is wound uniformly with an endless solenoid which produces an approximately uniform magnetising field $\frac{4\pi n_1 I}{10}$ (p. 232), where the current I in it is measured by the ammeter A.

A secondary coil of n_2 turns (Fig. 260) is also wound upon the ring, and is cut by the magnetic flux Ba, on establishing the magnetising field, B being the magnetic induction in the material, and a the area of cross-section of the ring. In series with the

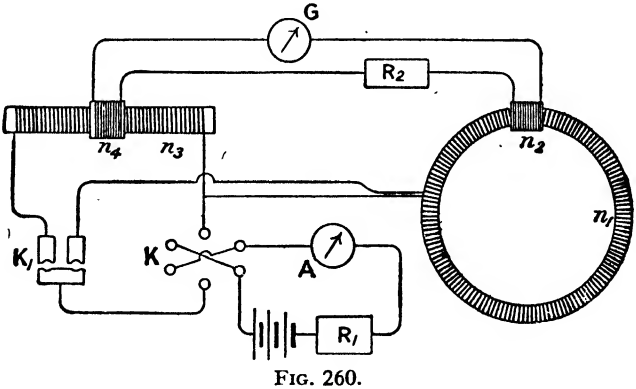


FIG. 260.

secondary coil is a ballistic galvanometer G. The charge caused to circulate in the galvanometer is then $\frac{Ban_2}{R}$, where R is the resistance of the secondary circuit, and if θ be the galvanometer throw, $\frac{Ban_2}{R}=k\theta$. k is determined by means of the standard

flux produced by the straight solenoid n_3 and its secondary n_4 (p. 265), and if θ_1 be the throw produced by establishing current I_1 in this standardising solenoid—

$$\frac{4\pi n_3 I_1 A n_4}{10R}=k\theta_1,$$

The two secondaries n_2 and n_4 being permanently in series with the galvanometer, R is the same in both cases, and therefore, eliminating k from the last two equations, we have—

$$B=\frac{4\pi n_3 n_4 A I_1}{10 a n_2 \theta_1}\theta.$$

From this and the equation—

$$H=\frac{4\pi n_1}{10}I,$$

the ballistic throws θ, and the ammeter readings I, may be converted into the corresponding values of B and H.

From an inspection of Fig. 261 it will be seen that on establishing a field Oh, and then removing it, the reverse throw on removal will be less than the direct throw on producing it, since a large fraction of the magnetisation remains; in fact, the throw obtained on establishing or removing the field depends so much on the previous condition of the specimen that it does not afford any useful information regarding its condition.

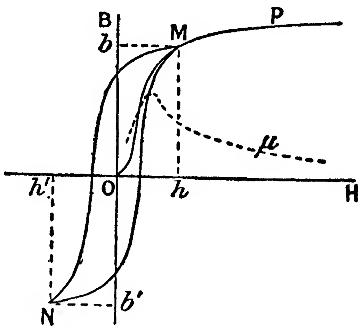


FIG. 261.

A better method of procedure is to obtain the throw for a reversal of the field, which of course gives 2B instead of B, for any value of H. Thus in Fig. 261, if the value of the magnetising field is Oh, and this be reversed to Oh'=—Oh, B changes from Ob to Ob', that is, the ballistic throw is proportional to bb'=2Ob. Instead of halving the value obtained for B to get that corresponding to the magnetising field Oh, it is usual to use a reversal of the current I_1 when obtaining the standardising throw θ_1, and since the throw for a reversal is made in both cases, the numerator and denominator in the quantity $\frac{4\pi n_3 n_4 A I_1 \theta}{10 a n_2 \theta_1}$ are both halved, which of course leaves it unaltered, and the fact of reversal may be ignored.

The method of procedure then is to apply the greatest magne-

tising current that is going to be used, and adjust the resistance R_2 until the ballistic throw on reversing the current by means of the key K (Fig. 260) makes full use of the galvanometer scale. This resistance must then remain constant throughout the observations. A few reversals of the current bring the iron into a steady cyclic condition, and then the throw θ for the reversal of the current I from positive to negative and again from negative to positive is observed, and the mean value of the two taken. The current is then reduced by means of the rheostat R_1, the repeated reversals performed and the process repeated. This is continued down to the smallest currents that give a reasonable throw. The θ's are then converted into B's, and the I's into H's, and the results plotted in the form of a curve. In Fig. 261 the curve OMP is obtained in this way for a ring of soft iron, and the permeability μ is calculated for each value of H by dividing B by H.

In this method there are no free polar surfaces, the tubes of magnetic induction being complete circuits within the iron; there is therefore no demagnetising effect, which is one of the advantages of this over the magnetometric method. Another advantage is that the galvanometer, if of the suspended coil type, is much less sensitive to outside magnetic disturbances than the magnetometer needle. Thus the employment of the magnetometric method requires the best laboratory conditions for success, but the ballistic method can be carried out almost anywhere. On the other hand, the ballistic method requires the welding, turning to circular form, and separate winding of each specimen examined, whereas in the magnetometric method any piece of the wire to be tested can be immediately placed in the magnetising solenoid for experiment. The ballistic method does not, as a rule, give us the cycle of magnetisation, but only the curve passing through the tips of the cycles for various magnetising fields.

Cycle by Ballistic Method.—If a cyclic curve of magnetisation be required, it may be found by a method devised by Prof. Ernest Wilson.[1] The key K, Fig. 260, is modified by replacing one of the cross conductors *ab* by a rheostat R_3 (Fig. 262). In making the reversals to establish the steady cycle of magnetisation, the short-circuiting tapping key T can be closed. To make the first measurement, T is kept closed and the rocker which connects *e* to *d* and *f* to *b* thrown over to *a* and *c*, which simply reverses the current. The reversal of H from the value OE to OF (Fig. 263) causes a throw proportional to the change of induction AB, and this is plotted downwards from A, the point C on the curve being obtained. The process is now repeated

[1] J. Hopkinson, E. Wilson, and F. Lydall, Proc. Roy. Soc., A 53 (1893), p. 352.

with T open, so that the magnetising field OE is reversed to the value OG, the corresponding throw being proportional to AL. This gives us the point M on the curve. T is then closed, the reversals made to re-establish the cycle, R_3 increased, T opened, and another throw obtained, giving the point P on the curve. The point S is found by merely breaking the circuit to reduce the field to zero, the throw proportional to AR being obtained. Points such as T are obtained by diminishing the field OE or suddenly increasing, by a small amount, the resistance of R_1

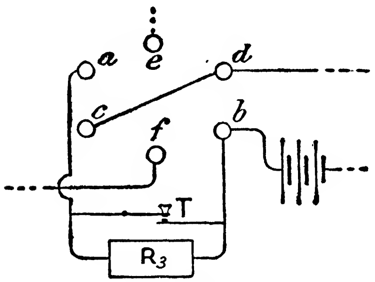


FIG. 262.

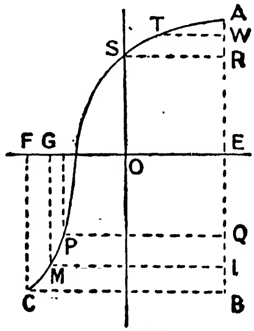


FIG. 263.

(Fig. 260). The other half of the cycle may then be drawn from symmetry.

The values of I may be obtained from those of B by means of the relation $B = H + 4\pi I$. Thus the I—H curve may be derived from the B—H curve, and *vice versâ*.

Cycle of Magnetisation—Hysteresis.—The behaviour of magnetic materials when subjected to cyclic changes of magnetic field, were first systematically studied by Sir J. A. Ewing, and for a detailed study of such cyclic changes, the student is referred to Ewing's work on "Magnetic Induction in Iron and other Metals." A typical cycle is seen in Fig. 264, the behaviour of the material being represented by the curve OABCDEFG. It will be observed that the descending branch of the curve always lies above the ascending branch. Hence the zero value of I occurs at a later point of the cycle than the zero value of H. To this lag of the magnetisation behind the magnetising field Ewing gave the name of *Hysteresis*.

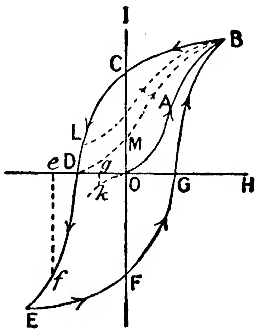


FIG. 264.

The value OC of the intensity of magnetisation when the

magnetising field is reduced from great values down to zero is called the *Residual Magnetism*. The value OD of the reversed field required to reduce the intensity of magnetisation to zero, is called the *Coercive Force*. A knowledge of the intensity of magnetisation near saturation, together with the values of the residual magnetism and the coercive force, enable one to draw approximately the magnetic cycle, and hence, failing the complete diagram of the cycle, these three quantities give a very good knowledge of the magnetic properties of the material.

The condition of the material at a point in the cycle represented by D, is very different from that of the neutral or unmagnetised condition represented by O; for a further negative field D*e* produces a large increase in the intensity of magnetisation represented by *ef*; while an equal negative field O*g*, applied to the unmagnetised substance, would only produce the small intensity *gk*. Or again, if, when the point L on the cycle is reached, the magnetising field, instead of being continued in the negative manner is brought back to its positive maximum, the dotted curve LB is followed, or if the return is made on reaching the point D, the path is the dotted curve DMB, in either case a closed loop being formed. Thus if the field is merely removed after the point D has been reached, there will be remaining magnetisation of intensity OM, and the specimen is certainly not demagnetised. The only satisfactory way to demagnetise a specimen is to take it repeatedly through cycles of continually decreasing range, ending with extremely small cycles; for the effect of one or two reversals of the field is to wipe out the effect of previous cycles, provided that there is not a great difference in range between the cycles. A specimen of iron may be demagnetised by heating it to red heat, and allowing it to cool in a region of no magnetic field; but this method is unsatisfactory, as the heating and cooling change the physical properties of the material.

Iron and Steel.—In Fig. 265 the curves are taken from Ewing's results. A is for annealed soft-iron wire, and B for the same wire after being hardened by stretching. C is for annealed pianoforte steel wire, and D for the same wire, glass-hard. We can see that the harder the material, the less is the residual magnetism, and the greater the coercive force. In Fig. 266 we have the curve E for annealed nickel wire and F when hardened by stretching. G is that for cobalt (containing 2 per cent. of iron). The curve for nickel resembles that for soft iron, but the saturation value of B is only about one-third of that for iron. The cobalt curve resembles that for steel, but the ascending and descending branches lie closer together. The saturation value of B for

cobalt is very little short of that for iron and steel. The coercive force for nickel is about 7·5 and for cobalt 12.

The effect of mechanical disturbance such as tapping is to make the ascending and descending branches for soft iron very nearly coincide; the residual magnetism and coercive force are

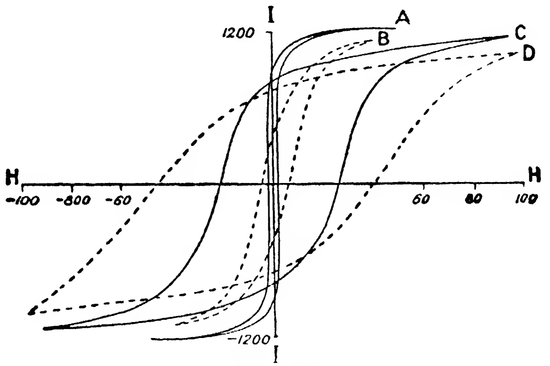


FIG. 265.

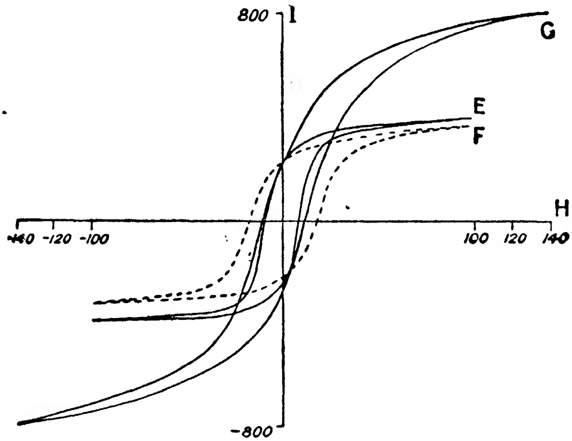


FIG. 266.

practically zero. The effect upon steel is in the same direction, but is not so marked.

Work due to Hysteresis.—The act of taking a body through a cycle of magnetisation involves the expenditure of energy; for the energy required to magnetise a specimen is not recoverable on removing the magnetic field, since the magnetisation does not fall to nothing; a negative field has to be applied before the intensity of magnetisation is brought to zero.

We will show from first principles that the work necessary to

produce a change dI in the intensity of magnetisation is HdI, dI being so small that the magnetising field H may be considered to be constant throughout the change.

Let m be the magnetic moment of one of the elementary magnets of the material, and θ the angle its axis makes with the direction of magnetisation.

$m \cos \theta$ and $m \sin \theta$ are the components of its moment parallel and normal to the direction of magnetisation. Then for the whole of the elementary magnets throughout unit volume, $\Sigma m \cos \theta = I$, the total magnetic moment per unit volume, and further $\Sigma m \sin \theta = 0$, otherwise there would be a component of the magnetic moment at right angles to the direction of magnetisation, which is contrary to the very meaning of the term.

From the former equation we have—

$$d\Sigma m \cos \theta = dI$$

i.e.

$$-\Sigma m \sin \theta \, . \, d\theta = dI.$$

Now the couple acting on the molecule m in the field H is

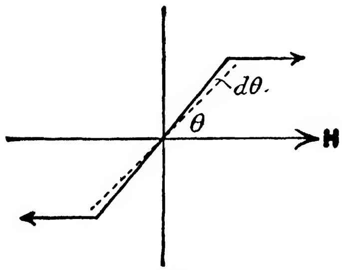


FIG. 267.

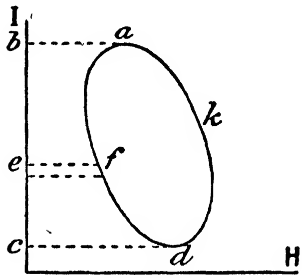


FIG. 268.

$mH \sin \theta$ (Fig. 267) ; and for a small rotation $-d\theta$ the work done is $-mH \sin \theta \, . \, d\theta$.

For all the elements in unit volume—

$$\text{work done} = -\Sigma m H \sin \theta \, . \, d\theta,$$
$$= -H \Sigma m \sin \theta \, . \, d\theta,$$

since H is constant. But this means an increase dI in the intensity of magnetisation, and

$$-\Sigma m \sin \theta \, . \, d\theta = dI,$$
$$\therefore \text{work done} = HdI.$$

Thus in the I—H diagram, Fig. 268, the work done for the small change dI in the intensity of magnetisation is HdI, that is, the area of the strip *ef*, and that in passing round the curve from *a* to *d* is the area of all such strips, that is the area *abecdfa*. Similarly the work done for the path *dka* is represented by the

area *dkabcd*; and the balance of work done on the specimen of unit volume, for the whole cycle is area *afdka*; thus $\int_0 \mathrm{H}d\mathrm{I}$ is the work done for any cyclical path, where $\int_0$ is the integral round the whole path. We can therefore see that the work for a cycle of magnetisation of steel is much greater than that for soft iron since the hysteresis curve encloses a much larger area. This work appears in the form of heat in the specimen and represents an irrecoverable loss of energy. Not only is energy lost, but the heating effect is cumulative, and in a mass of iron subjected to an alternating magnetic field, as in the armature core of a dynamo, or the core of a transformer, the consequent rise of temperature may be considerable. For this reason the iron used for these purposes has as low a hysteresis effect as possible.

The area $\int_0 \mathrm{H}d\mathrm{I}$ may be obtained from any of the I—H cycles, paying due regard to the scale upon which the curve is drawn. If the curve be one for B and H, the area must be divided by 4π to obtain the work done per cubic centimetre per cycle.

For, $$\mathrm{B}=\mathrm{H}+4\pi\mathrm{I},$$

$$\therefore\ \mathrm{H}d\mathrm{B}=\mathrm{H}d\mathrm{H}+4\pi\mathrm{H}d\mathrm{I},$$

and, $$\int_0 \mathrm{H}d\mathrm{B}=\int_0 \mathrm{H}d\mathrm{H}+4\pi\int_0 \mathrm{H}d\mathrm{I}.$$

The term $\int_0 \mathrm{H}d\mathrm{H}$ is necessarily zero, for if we plot H against H, we get a straight line, and the area enclosed for any cycle will of course be zero.

The value of $\int_0 \mathrm{H}d\mathrm{I}$ varies from about 10,000 ergs for annealed soft iron to 117000 ergs for hardened pianoforte steel wire.

Taking the density of iron as 7·7 and its specific heat 0·11, the thermal capacity of 1 c.c. is $7{\cdot}7\times0{\cdot}11$, and the rise in temperature per cycle of magnetisation is $\dfrac{\int_0 \mathrm{H}d\mathrm{I}}{7{\cdot}7\times0{\cdot}11\times4{\cdot}2\times10^7}$ degrees; the mechanical equivalent of one calorie being $4{\cdot}2\times10^7$ ergs. For a value of 50,000 for $\int_0 \mathrm{H}d\mathrm{I}$, and a frequency of 100 cycles per second, we have a rise of temperature per second of—

$$\frac{50000\times100}{7{\cdot}7\times0{\cdot}11\times4{\cdot}2\times10^7}=0{\cdot}14^\circ,$$

or 8·4° per minute, provided that the heat produced did not leak away.

Hysteresis Tester.[1]—The importance of this hysteresis loss of energy in a magnetic material led Sir J. A. Ewing to devise a piece of apparatus by means of which the hysteresis loss in a specimen of the material may be found, without making the laborious test for finding the magnetic cycle. The specimen is rapidly rotated between the poles of a permanent magnet, which

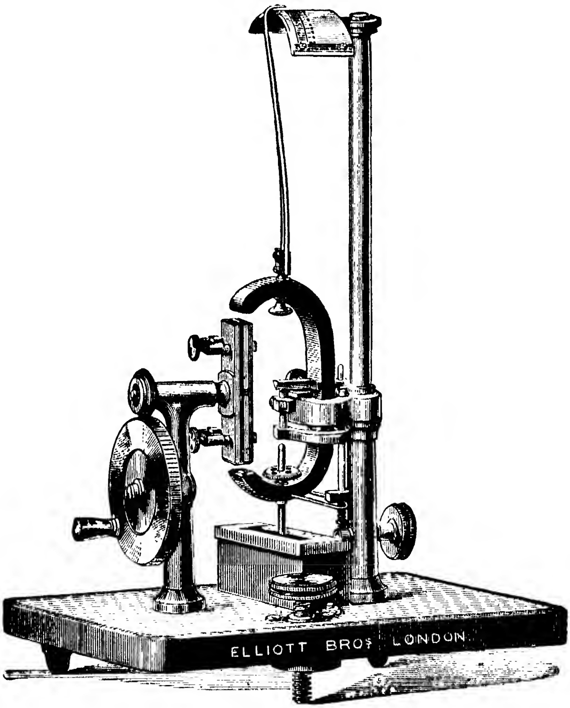


FIG. 269.

is supported upon knife-edges to enable it to turn about a horizontal axis (Fig. 269).

As the specimen rotates, it is magnetised by the field of the permanent magnet, the lag in polarity causing it, by the attraction between the respective poles, to drag the magnet after it. The deflection of the magnet is measured by the pointer and scale, and is proportional to the hysteresis effect in the specimen, being independent of the speed of rotation. This may be shown in the following manner :—

Let H be the field due to the permanent magnet at any point, and θ the angle between H and the direction of magnetisation of

[1] J. A. Ewing, Inst. Elec. Eng., vol. 24, p. 398. 1895.

the specimen at the point. As the specimen makes one complete rotation, its magnetisation at the point considered goes through a cycle, and the work done due to hysteresis is $\int_0 \mathrm{H}d\mathrm{I}$ per unit volume.

Thus, at the point considered, H in this expression must be replaced by H cos θ, and further, $\mathrm{I}=\kappa\mathrm{H}\cos\theta$, where κ is the susceptibility.

$$\therefore\ d\mathrm{I}=-\kappa\mathrm{H}\sin\theta\,.\,d\theta.$$

Therefore the work done for one cycle is—

$$\int_0 \mathrm{H}\cos\theta(-\kappa\mathrm{H}\sin\theta)d\theta=-\int_0 \kappa\mathrm{H}^2\sin\theta\,.\cos\theta\,.\,d\theta.$$

Again, to find the couple exerted on the specimen,—I is the magnetic moment of unit volume, and the couple on it is therefore IH sin θ, tending to rotate the specimen into the direction of H, and—

$$\mathrm{I}=\kappa\mathrm{H}\cos\theta,$$

$$\therefore\ \text{couple}=\kappa\mathrm{H}^2\sin\theta\,.\cos\theta.$$

The mean value of this for a complete rotation is—

$$\frac{\int_0 \kappa\mathrm{H}^2\sin\theta\,.\cos\theta\,.\,d\theta}{\int_0 d\theta}=\frac{1}{2\pi}\int_0 \kappa\mathrm{H}^2\sin\theta\,.\cos\theta\,.\,d\theta.$$

Comparing this with the expression for the work done due to hysteresis, we see that the variable parts are identical, and therefore the mean couple acting between the specimen and the permanent magnet is proportional to the hysteresis effect, and is independent of the speed of rotation. It is balanced by the gravitational couple, which is measured by the deflection of the permanent magnet from its mean position.

The instrument is calibrated by means of two specimens, one of low and the other of high hysteresis value, and the samples to be tested are made of the same size and shape as the standards, the length being the important quantity to have correct.

Steinmetz Law.—An empirical formula for the work done in a cycle of magnetisation has been given by Steinmetz,[1] which is very useful for many practical purposes ; it states that the work per cycle is proportional to the magnetic induction raised to a constant power which ranges between 1·66 and 1·70.

Thus—
$$\int_0 \mathrm{H}d\mathrm{I}=\eta\mathrm{B}^{1\cdot 68},$$

[1] C. P. Steinmetz, *Electrician*, **26**, p. 261 (1891) ; **28**, p. 425 (1892).

where η is a coefficient depending upon the material, and B the maximum value of the induction during the cycle.

For	very soft iron	$\eta=0{\cdot}0020$.
,,	hardened steel	$\eta=0{\cdot}025$.
,,	annealed cast-steel	$\eta=0{\cdot}0080$.
,,	nickel	$\eta=0{\cdot}012$ to $0{\cdot}038$.
,,	cobalt	$\eta=0{\cdot}012$.

The law only roughly represents the truth, and can only be used for approximate purposes.

Iron and Steel Alloys.—Many substances, such as silicon, chromium, tungsten and manganese, in small quantities, profoundly modify the magnetic properties of steel. Thus chromium, tungsten or manganese, in small quantities (up to 4 per cent.), greatly increase the coercive force, in some cases up to 40 or even 50, while 12 per cent. of manganese (Hadfield's manganese-steel) renders it almost non-magnetic at low fields, the permeability being about 1·4 for all fields. Also mumetal (Ni 76, Fe 17, Cr 1·5 and Cu 5 per cent.) has extremely high permeability and on this account has many uses, as has radiometal (Ni 48, Fe 48, Cu 3, Mn 0·5). Mumetal has a very high permeability (10,000 to 100,000) and low hysteresis (60 ergs per cycle per c.c. for $B=5000$) and the corresponding values for radiometal are 2000 to 15,000 and 350.

Magnetic Alloys of Non-magnetic Substances.—It was found by Heusler [1] that it was possible to produce a magnetic alloy of non-magnetic substances. Thus several alloys of manganese, aluminium and copper, and of manganese, aluminium and zinc, exhibit marked magnetic properties. An alloy of 26·5 per cent. Mn, 14·6 Al and 58·9 Cu, has a permeability of 225 for a magnetising field of strength 20.[1] The magnetic behaviour of these alloys depends very much upon their previous condition with regard to temperature.

Force between Magnets in Contact.—When two magnetic polar faces are in contact, as in the case of a soft iron core divided transversely, there is a force pulling the two polar faces together. Let σ be the amount of pole per unit area of face, N on one side and S on the other. Then each produces a field of strength $2\pi\sigma$, and the other polar face being situated in this, experiences a force $2\pi\sigma^2=F$ per unit area. The two faces are not in contact at more than a few points, and the strength of field H in the air interspace is $4\pi\sigma=H$ (p. 126).

$$\therefore\ F=\frac{H^2}{8\pi}.$$

[1] Fr. Heusler, "Verh. D. Phys. Ges.," 1903.

But the magnetic field being normal to the faces, the value of H in the gap is equal to the value of B in the iron (p. 234).

$$\therefore F = \frac{B^2}{8\pi}.$$

Weak Magnetic Fields.—The experiments described above are not sufficiently delicate to determine the form of the curve of magnetisation very near the origin. The late Lord Rayleigh [1] examined this point, and found that for very weak fields κ and μ are constant, and hence the I—H and B—H curves are practically straight lines near the origin, and are inclined to the axis of H. The method he adopted was to place the specimen inside a magnetising coil B (Fig. 270), with its end very close to the magnetometer needle. For moderate field strength the effect on the magnetometer needle is balanced by a coil A, in series with B. On varying the current, the balance is still perfect if the permeability is everywhere constant, but if that of the specimen varies, the effect of A and B on the needle will not change at the same rate, and the balance is destroyed. With a piece of Swedish iron wire the balancing was made with a field of strength 0·04, and it was found to remain perfect as the field was reduced to 0·00004 Hence, for these fields the permeability is constant. Up to values of H equal to 1·2, the following formulæ gave, fairly well, the values of μ and κ.

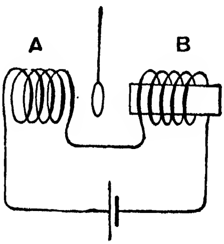


FIG. 270.

$$\kappa = 6{\cdot}4 + 5{\cdot}1H.$$
$$\mu = 81 + 64H.$$

Time Lag.—Sir J. A. Ewing found that in the case of soft iron, the specimen did not take its final value of the magnetisation instantaneously, and in employing the magnetometer, an interval had to be allowed to elapse before reading the deflection, to allow the magnetisation to creep up to its full value. This renders the readings taken by the ballistic method somewhat uncertain, but if the ballistic galvanometer were replaced by the Grassot Fluxmeter (p. 266) this difficulty would be removed. With hard iron and steel, Lord Rayleigh found that there was no time lag for weak fields. Using a method similar to Lord Rayleigh's, Ewing [2] found that annealed wrought iron took in some cases as long as 60 seconds to creep up to its final magnetisation for fields not exceeding 0·1, but the greater part of the magnetisation was acquired within 5 seconds. A cycle of

[1] Lord Rayleigh, *Phil. Mag.*, **23**, p. 225. 1887.
[2] J. A. Ewing, *Proc. Roy. Soc.*, **46**, p. 269. 1889.

magnetisation may therefore be produced in a variety of ways. Thus for the weak field Oh (Fig. 271) suddenly applied, the resulting intensity of magnetisation is ha, but after a time this creeps up to hb. If the field had been applied very slowly, the path Ob would have been followed. On then suddenly reversing the field bd' is the curve followed, and with time the point b' is slowly reached. Another reversal, followed by a pause, gives the path $b'db$. Thus for a cycle consisting of rapid reversals, the magnetisation curve is $aoa'oa$, and there is no hysteresis loss; for very slow change of field the curve is $bob'ob$, again with no hysteresis; but for any other change there is always a loop and consequently hysteresis loss, reaching a maximum for the path $bd'b'db$.

Very Strong Fields.—In order to determine whether the intensity of magnetisation really approaches a limiting saturation value, as indicated by the molecular theory, Prof. Ewing and

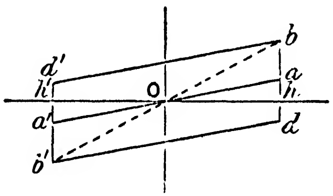


Fig. 271.

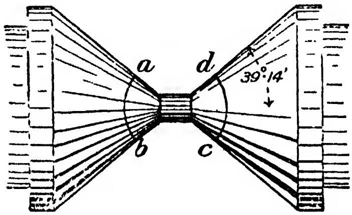


Fig. 272.

Mr. Low [1] employed what they called the Isthmus method. The specimen forms a neck or isthmus between the tips of the conical poles of an electromagnet. They showed that for greatest uniformity of field at the neck, the semi-angle of the cone should be 39° 14′ while for greatest value of the field it should be 54° 44′. Both forms were used. The magnetic induction in the specimen forming the neck was measured by winding a coil on it and rotating it through 180°, the throw of a ballistic galvanometer in series with the coil being observed. In order to make the rotation possible, the tips of the pole pieces are bored through transversely by a circular hole $abcd$ (Fig. 272), and an iron bobbin with the neck as shown placed to fill the hole. The strength of magnetising field is determined by winding a second coil outside the first, so that it encloses an air space of known section. The difference in the ballistic throws for the two coils is proportional to the magnetic flux through the air space between the coils, and therefore to the magnetising field.

[1] J. A. Ewing and W. Low, *Phil. Trans.*, A., **180** (i), p. 221. 1889.

The results for a specimen of Vicker's tool steel are given below—

H	B	I	μ
6210	25480	1530	4·10
9970	29650	1570	2·97
12120	31620	1550	2·60
14660	34550	1580	2·36
15530	35820	1610	2·31

It will be seen that the intensity of magnetisation has become very nearly constant, and μ seems to be approaching the value unity, which it should have for infinite fields if I ceases to increase ; for we see from the expression—

$$B = H + 4\pi I,$$

that if H becomes very great compared with $4\pi I$, the latter is negligible, and $B = H$, or $\mu = 1$.

Much stronger fields have been obtained by Kapitza [1] by short circuiting an alternating current generator through the coil which produces the magnetising field. The current lasts for about $\frac{1}{100}$ second but produces a magnetising field of about 300,000 gauss. The current is measured by means of an oscillograph and the magnetic field by means of a search coil which is short circuited until the current has flowed for about half its time when the short circuit is broken. It was seen on p. 236 that the force on a small body in a field which is not uniform is $\frac{v\kappa H dH}{ds}$. Now κ is the volume susceptibility, whereas the mass susceptibility is χ, or the magnetic moment of unit mass for unit magnetising field. Then $\chi = \frac{\kappa}{\text{density}}$, $\kappa = \chi \times \text{density}$, and $v\kappa = \chi(\text{density} \times \text{volume}) = \text{mass} \times \chi$. So that force$= m\chi H\frac{dH}{ds}$ $= \frac{m}{2}\chi\frac{dH^2}{ds}$. It will be seen that if χ is constant the force is proportional to $\frac{dH^2}{ds}$. Kapitza found this to be true for paramagnetic and diamagnetic substances. Whereas for ferromagnetics the force is proportional to $\frac{dH}{ds}$, which indicates that the magnetisation χH is constant. It appears then that ferromagnetic saturation has been attained in these strong fields.

[1] P. Kapitza, *Proc. Roy. Soc.*, **131, 1931.** Also see " Magnetism and Matter," by Edmund C. Stoner.

Variation of Temperature.—All magnetic materials vary in susceptibility when the temperature changes. Hopkinson [1] found that in general, the susceptibility increases with rising temperature when the specimen is subjected to a weak magnetising field, but for strong fields the reverse is the case. At a dull red heat, iron loses its magnetic properties entirely, the temperature at which this occurs varying from 690° C. to 870° C. for different materials. The change does not take place suddenly, but in a few degrees rise in temperature the iron changes from a highly magnetic to a non-magnetic substance. This critical temperature is also the temperature at which recalescence or the sudden reglowing of a mass of cooling iron occurs, and it was shown by Tait that the thermo-electric power of the magnetic metals undergoes rapid changes at the critical temperature. Evidently some important molecular rearrangement occurs at this temperature, and that this arrangement is intimately asso-

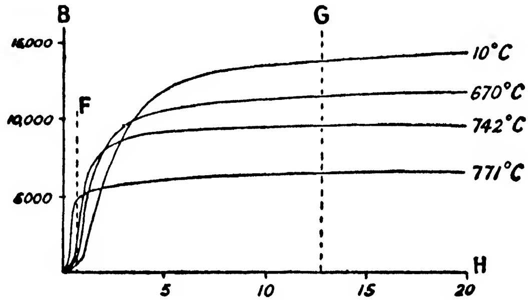


Fig. 273.

ciated with the acquirement of magnetic properties on cooling, is shown by the fact that non-magnetisable manganese steel (12 per cent. Mn, 1 per cent. C) does not exhibit the phenomenon of recalescence. From Fig. 273, taken from Hopkinson's results, it is seen that in the case of soft wrought-iron, for fields below 0·5 the susceptibility increases with temperature, while for strong fields the susceptibility falls with rising temperature, as will be seen on comparing the changes of B with temperature along the vertical lines F and G. At a temperature of 788° the material has become non-magnetic.

If the permeability be plotted against the temperature for three fields 0·3, 4 and 45, the diagram, Fig. 274, is obtained. It will be seen that for low magnetising fields the permeability increases rapidly as the critical temperature is reached, but for high fields in which the value of B is of course much nearer saturation value throughout, the permeability is not so much

[1] J. Hopkinson, *Phil. Trans.*, A., **180** (i), p. 443. 1889.

affected by temperature, and in all cases the permeability becomes zero at 785° C.

For nickel, Hopkinson [1] found the critical temperature to be 310° C. and the general course of the phenomenon similar to that in iron.

Ewing has shown [2] that there is no temperature hysteresis unless the range of variation of temperature includes the critical temperature. That is, on heating and cooling the metal in a magnetic field the intensity of magnetisation at any temperature is the same when the temperature is falling as when it is rising.

It has been found by Curie [3] that besides the great change that takes place at the critical temperature, there are others at still higher temperatures, the most important of which consists in a

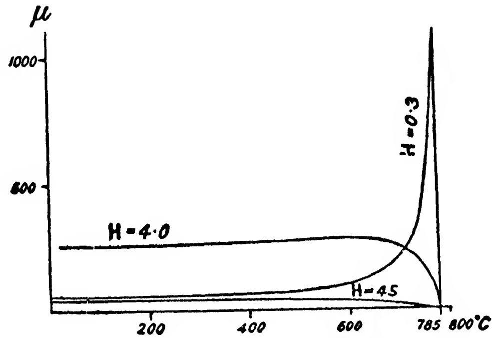


FIG. 274.

sudden rise in magnetisation at 1280° C. On plotting the magnetisation-temperature curve to a very large scale, it is seen that between 750° and 800° C. the intensity of magnetisation drops to one-hundredth of its value, and then continues to decrease, with a slight inflection at 860° C. At 1280° C. I suddenly rises from about 0·025 to 0·040, and from then again decreases.

Mechanical Stress.—The effects of mechanical stress upon the magnetic properties of materials are exceedingly complicated, but the following are the most striking.

It was found by Villari [4] that for weak fields, longitudinal tension increases the magnetisation, but for strong fields the reverse is the case. Thus the curves for annealed soft iron wire subjected to a pull are as shown in Fig. 275, after Ewing.[5] Also

[1] J. Hopkinson, *Proc. Roy. Soc.*, **44**, p. 317. 1888.
[2] J. A. Ewing, *Phil. Trans.*, **176**, p. 523. 1885.
[3] P. Curie, *Comptes Rendus*, **118**, 726, 859, 1134.
[4] E. Villari, *Pogg. Ann.*, p. 322. 1868.
[5] J. A. Ewing, *loc. cit.*

the effect of varying the load in a cyclic manner, when the field is constant, is similar in character at all fields, but varies in amount with the field. With a magnetising field of 0·34 the

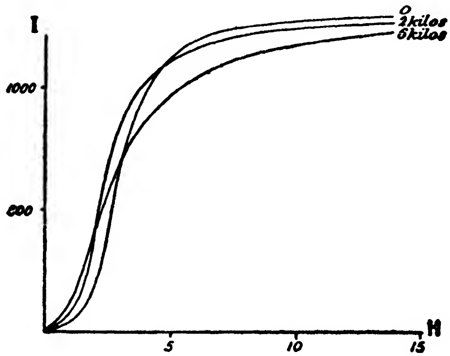


FIG. 275.

effect of increasing and then decreasing the load is seen in the lower dotted curve (Fig. 276), and on again increasing and decreasing the load, the curve becomes cyclic, as shown by the

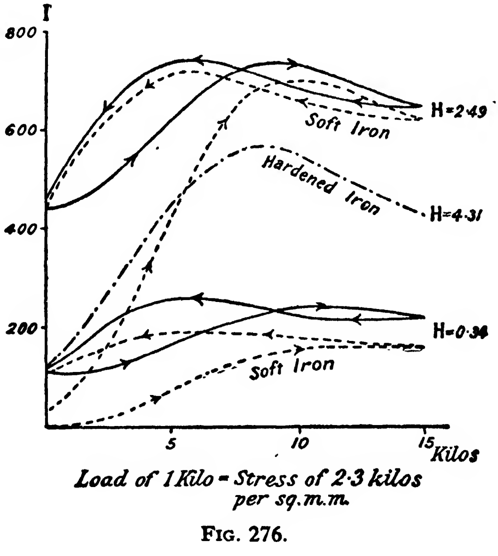


FIG. 276.

full line. With a magnetising field of 2·49 the effects are similar and are much more marked. Fewer reversals are necessary to reach the cyclic state at high magnetising fields than at low fields. There is no hysteresis for steel and hard iron, for which the chain curve is typical. Nickel shows no Villari reversal,

magnetisation being decreased by longitudinal tension and increased by compression at all field strengths.[1]

Lord Kelvin, who also investigated the effect of stress on magnetisation, showed[2] that a stress applied transversely has an opposed effect to a similar longitudinal stress. As might be expected, torsion gives many interesting and complicated effects.

Magneto-striction.—Changes in dimensions with magnetisation are called magneto-strictive effects. Shelford Bidwell[3] used an optical lever system measuring changes in length to one part in 10^7, Fig. 277 representing some of his results. If magnetisation occurs while the specimen is under stress, the curve changes and in the case of iron deforms continuously as indicated by the dotted curves.

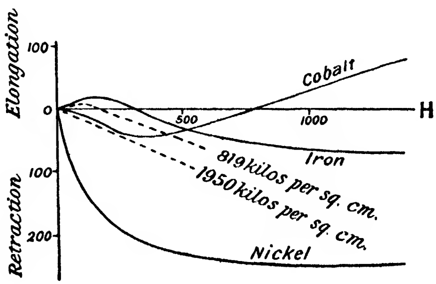


Fig. 277.

The magneto-striction in paramagnetics and diamagnetics is so small with ordinary fields that it is immeasurable. With fields up to 300,000 gauss, Kapitza[4] has measured it for a number of such substances and finds that the relative change in length is proportional to the square of the applied field strength.

Magneto-strictive oscillators.—A specimen of, say, nickel can be forced into resonant vibration by an alternating magnetic field adjusted in frequency to match a natural frequency of mechanical vibration of the specimen. If such a specimen is immersed in a liquid, powerful ultrasonic waves may be generated. Conversely, mechanical vibrations induced by receipt of ultrasonic waves change the magnetisation of such a specimen synchronously and may be made to record by using the induced alternating E.M.F. in a secondary coil wound on the specimen. The interval between the generation and the receipt of pulses of ultrasonic waves reflected from the ocean floor is used in depth sounders.[5]

Molecular Theory.—Many of the characteristic features of magnetic substances can be explained by a " molecular theory " of some form, the first promising theory of this kind being that proposed by W. Weber (1852). The magnetic properties exhibited by a body are attributed to an ordered arrangement of minute permanent magnets, the " molecules " of magnetism ; if

[1] J. A. Ewing, *Phil. Trans.*, A., **179** (i), pp. 325 and 333. 1888.
[2] Sir W. Thomson, *Proc. Roy. Soc.*, **27**, p. 439. 1878.
[3] Shelford Bidwell, *Proc. Roy. Soc.*, pp. 109, 257 (1886) ; *Phil. Trans.*, p. 469 (1890) ; **179** (i), p. 205 (1888).
[4] P. Kapitza, *Proc. Roy. Soc.*, **135**, 1932.
[5] A. B. Wood, F. D. Smith and J. A. McGeachy, *Journ. Inst. Elec. Eng.*, **76**, p. 550. 1935.

these magnets are in a purely random arrangement, the specimen is not magnetised.

It is supposed that in a substance which is uniformly magnetised to saturation, the molecular magnets are all aligned. At a sufficiently large distance from a bar so magnetised, there will be no magnetic effect due to the elementary magnets within the bulk of the substance because of the cancelling effects of the equal numbers of N and S poles. Thus polarity is only evident at the ends of the specimen. The absence of translatory force on a magnet in a uniform field (demonstrated by a magnet on a cork floating in water) is referred back to the assumed equality of poles in each elementary magnet or dipole. The appearance of fresh poles on the new faces if the material is broken transversely is also readily explained.

If the rotation of the " molecules " into alignment with a superposed field is opposed by some form of frictional constraint, some of the features of the typical hysteresis curve (p. 279) appear, but something akin to elastic constraint is required to account for

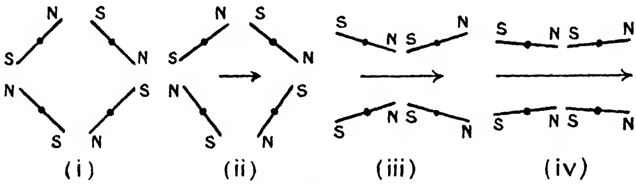


FIG. 278.

the reversible magnetisation which, as Lord Rayleigh showed (p. 287), occurs in very weak fields.

Sir J. A. Ewing showed that mechanical constraints need not be invoked since the phenomena of hysteresis could be explained by magnetic interactions between the particles, provided these forces were sufficiently strong. This theory will be found fully expounded in Ewing's book, " Magnetic Induction in Iron and other Metals." The essential features may be understood with the help of Fig. 278, which represents the successive stages in the magnetisation of a group of four " molecules," giving rise to the initial magnetisation curve shown in Fig. 279. If a vast number of such groups, of varying stability, are present in a specimen, the overall effect of contributions from all groups will be to produce a smooth curve such as those in Figs. 261 and 273. The theory is impressively supported by watching the behaviour of an array of pocket-compasses (30 or more) subjected to a field (from a large coil carrying current or due to a group of bar magnets) which is increased, decreased and reversed.

The size and nature of the " molecule " and the true nature of the forces acting between the elements involved remained in doubt. Valuable evidence that magnetisation involved dis-

continuous steps was provided by the Barkhausen effect.[1] To demonstrate this phenomenon, a specimen S (Fig. 280), such as a rod of iron, is taken through a cycle of magnetisation by means of a coil M. A secondary coil C, of many turns, in series with a compensating coil K so that inductive effects from M cancel, is connected to a valve amplifier A (see Chapter XVI), the output of which is fed to a telephone or loudspeaker T. While the steep part of the magnetisation curve is being traversed, a random series of impulses is induced in C, giving rise to clicks in the telephone which may merge into a murmuring or " rushing " sound. These impulses may be examined and recorded by using an Einthoven galvanometer (p. 75) connected in place of T.

Careful study of the Barkhausen effect has been made by many workers and it has been shown[2] that the inductive kicks are due to sudden changes in magnetisation of certain regions or

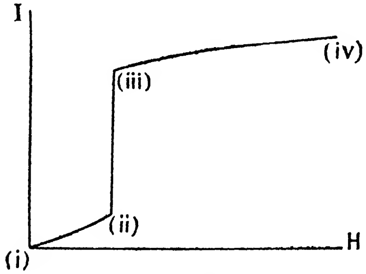


FIG. 279.

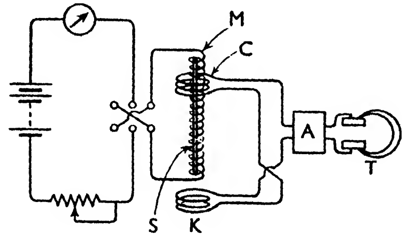


FIG. 280.

" domains " in the specimen, each containing on the average some 10^{15} atoms. Since, as we shall see later (Chapter XVI), interatomic distances in solids are of the order 10^{-8} cm., a compact group of this number will have linear dimensions of about $10^5 \times 10^{-8} = 10^{-3}$ cm., so that these domains are microscopic in size.

Direct experimental evidence of the existence of a domain structure was obtained by F. Bitter,[3] who examined under the microscope polished surfaces of crystals of iron, cobalt and nickel which had been coated with a suspension of very fine magnetic particles. This technique is used to detect cracks in ferromagnetic material which has been magnetised, since the particles are attracted (p. 237) to the intense local fields which arise at discontinuities. The *Bitter Figures* appear on unmagnetised as well as on magnetised material and although their interpretation is not always easy, they clearly demonstrate the existence of spontaneously magnetised domains. These domains are not necessarily interrupted by the boundaries between the close-packed crystals of which ordinary polycrystalline metals are composed.

Modern theory of the atom, discussed in Chapter XVI, suggests

[1] H. Barkhausen, *Zeit. für techn. Physik*, **5**, p. 518. 1924.
[2] R. M. Bozorth and Joy Dillinger, *Phys. Rev.*, **35**, p. 733. 1930.
[3] F. Bitter, *Phys. Rev.*, **38**, p. 1903 (1931) and **41**, p. 507 (1932).

that ferromagnetism is ultimately due to spinning electrons, or elementary electric charges. These act as magnetic dipoles and their alignment in magnetisation should involve the substance acquiring a spin so that angular momentum may be conserved.

Such a gyromagnetic effect can be observed[1] by optical recording of the small angular rotation shown by a rod of iron or other suitable material, suspended by a fine fibre so as to hang along the axis of a vertical solenoid. The ratio of the angular momentum acquired by the rod to its magnetic moment, the so-called *gyromagnetic ratio*, is the ratio for the individual magnetic elements, supposed all alike. For ferromagnetic substances this ratio is found equal to that expected for the spinning electron.

Resonance methods are also used,[2] the magnetic field being reversed in synchronism with the vibrations of the suspended specimen. It has even proved possible to work with paramagnetics.[3]

The Magnetic Circuit.—We have seen that magnetic flux is distributed circuitally, for at the surface of separation of two media the normal magnetic induction is continuous as it crosses the surface (p. 234), and in a uniform medium the tubes of induction are continuous. Consider a tube of induction whose cross-section at any point is s, the value of the induction at the section being B. The total magnetic flux over this cross-section is Bs, if B is uniform; and if B is not uniform then $\int$Bds is the flux. This quantity is constant for every section (p. 234) and is therefore characteristic of the tube.

Let us now find the line integral of the magnetic field for a circuital path round the tube of induction. It may be expressed in the form $\int_0 Hdl$, if H is everywhere parallel to the element of the path dl, which will be true if the tube is sufficiently narrow; but in any case $\int_0 H\cos\epsilon\, dl$ will be the line integral, where ϵ is the angle between the directions of H and dl. This line integral is the work done on carrying a unit pole once round the path, and by analogy with the corresponding electrical case, it is sometimes called the *Magneto-Motive Force* (M.M.F.) round the complete circuit formed by the tube of induction.

We have seen that the line integral of the magnetic field round any closed path is equal to 4π times the total current linked with the path.

$$\therefore \int_0 Hdl = \frac{4\pi nI}{10},$$

where I is the current in amperes flowing in a wire linked with

[1] J. Q. Stewart, *Phys. Rev.*, **11**, p. 100 (1918); A. P. Chattock and L. F. Bates, *Phil. Trans. Roy. Soc.*, A, **223**, p. 257 (1923).

[2] A. Einstein and W. de Haas, *Ver. der d. Phys. Ges.*, **17**, p. 152 (1915); G. Beck, *Ann. der Phys.*, **60**, p. 109 (1919).

[3] W. Sucksmith, *Proc. Roy. Soc.*, A, **128**, p. 276 (1930); **133**, p. 179 (1931); **135**, p. 276 (1932).

the magnet circuit, and n the number of times the two circuits are linked together. The product nl is frequently called the number of ampere-turns linked with the circuit, and hence we may now write for any magnetic circuit—

$$\text{M.M.F.}=\frac{4\pi}{10}\times(\text{ampere-turns}).$$

Again, for the circuital tube of induction—

$$\text{magnetic flux } \text{N}=\text{B}s=\mu\text{H}s$$

$$\therefore\ \text{H}=\frac{\text{N}}{\mu s},$$

and, $$\text{M.M.F.}=\int_0 \text{H}dl=\int_0\frac{\text{N}dl}{\mu s}=\text{N}\int_0\frac{dl}{\mu s}$$

since N is constant for the circuit.

$$\therefore\ \text{N}=\frac{\text{M.M.F.}}{\int_0\frac{dl}{\mu s}}.$$

By analogy with the case of an electric current circuit for which—

$$i=\frac{\text{E.M.F.}}{\int_0\frac{\text{S}dl}{s}}$$

where $\int_0\frac{\text{S}dl}{s}$ is the resistance, S being the resistivity at any point, the quantity $\int_0\frac{dl}{\mu s}$ is called the *magnetic resistance* or *reluctance* of the circuit. It must be remembered, however, that the resemblance is only in the form of the expressions, as there is no such thing as a magnetic current. The quantity N, although called a magnetic flux is only a statical condition defined by the relation $\text{N}=\int\mu\text{H}ds$.

When the shape of the tube of induction is completely known, and also the value of μ at each point, the quantity $\int_0\frac{dl}{\mu s}$ can be found. In certain simple cases the circuit may consist of several parts, for each of which μ and s are constant, and when this is so, the magnetic resistance of the whole circuit is the sum of the magnetic resistances of these separate parts. In many other cases, where the boundary of the circuit considered is not everywhere parallel to the direction of the flux, useful approximate values for the magnetic resistance may be obtained by following a similar method, but in this case the magnetic circuit is not perfect, and uncertainty in calculation is introduced by the uncertainty of the dimensions of the nearest perfect circuit.

The following examples will illustrate the method :—

(i) *Ring wound with Endless Solenoid.*—In this case the magnetic circuit consists of one homogeneous iron ring ; B, H and μ being approximately constant for all points. The magnetic resistance of the ring is $\int_0 \frac{dl}{\mu s} = \frac{2\pi r}{\mu s}$.

Hence if there are n turns per centimetre length of ring, total turns $= 2\pi rn$. With current I amperes in each turn—

$$\text{ampere turns} = 2\pi rn\text{I},$$

$$\text{and M.M.F.} = \frac{4\pi(2\pi rn\text{I})}{10};$$

$$\therefore \text{N} = \frac{\frac{4\pi(2\pi rn\text{I})}{10}}{\frac{2\pi r}{\mu s}} = \frac{4\pi n\text{I}\mu s}{10}.$$

But,
$$\text{B} = \frac{\text{N}}{s} = \frac{4\pi n\text{I}\mu}{10},$$

and,
$$\text{H} = \frac{\text{B}}{\mu} = \frac{4\pi n\text{I}}{10},$$

a result which we obtained previously on p. 232.

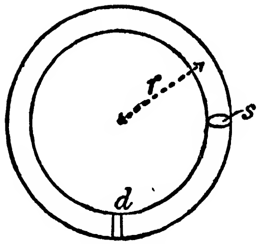


FIG. 281.

(ii) *Ring with Small Air-Gap.*—If the air-gap has thickness d (Fig. 281), the magnetic circuit consists of two parts, $(2\pi r - d)$ cm. of iron, having magnetic resistance $\frac{2\pi r - d}{\mu s}$, and d cm. of air, having magnetic resistance $\frac{d}{s}$, since $\mu = 1$.

$$\therefore \text{whole magnetic resistance} = \frac{2\pi r - d}{\mu s} + \frac{d}{s}$$

$$= \frac{2\pi r + (\mu - 1)d}{\mu s}.$$

Then,
$$\text{N} = \frac{4\pi(2\pi rn\text{I})}{10} \cdot \frac{\mu s}{\{2\pi r + (\mu - 1)d\}}.$$

By comparison with the value of N for the ring without gap, it will be seen that the magnetic effect of the gap is to increase apparently the length of iron in the ring by an amount $(\mu - 1)d$, which is approximately $1000d$, when $\mu = 1000$. Hence the enormous drop in magnetisation due to quite a small gap.

As before,
$$\text{B} = \frac{\text{N}}{s} = \frac{4\pi(2\pi rn\text{I})\mu}{10\{2\pi r + (\mu - 1)d\}},$$

and the actual value of H within the iron, where the permeability is μ, is—

$$H=\frac{B}{\mu}=\frac{4\pi(2\pi rnI)}{10\{2\pi r+(\mu-1)d\}}.$$

Calling the corresponding value of the magnetising field when there is no gap, H′—

$$H'=\frac{4\pi(2\pi rnI)}{10(2\pi r)},$$

and we have—

$$\frac{H'}{H}=\frac{2\pi r+(\mu-1)d}{2\pi r}=1+\frac{(\mu-1)}{2\pi}\cdot\frac{d}{r}.$$

But $\frac{d}{r}=\theta$, the angular thickness of the gap;

$$\therefore\ H'=H+\frac{(\mu-1)}{2\pi}H\theta.$$

Now, $\mu=1+4\pi\kappa \qquad \therefore\ \mu-1=4\pi\kappa,$

$$\kappa H=I,$$
$$H'=H+2I\theta$$
$$=H+\frac{4\pi\theta^\circ}{360}I$$

where θ° is given in degrees instead of radians, for—

$$\theta^\circ=\frac{360}{2\pi}\theta.$$

The equation is therefore $H=H'-\frac{4\pi\theta^\circ}{360}I$, and it will be seen by comparison with the equation on p. 271, that the gap produces a demagnetising effect, the coefficient of demagnetisation being $\frac{4\pi\theta^\circ}{360}$. When the gap has a thickness of half a degree—

$$H=H'-0{\cdot}0174I.$$

And it will be seen from the table on p. 273, that the demagnetising effect of the gap is nearly the same as that for an ellipsoid whose length is fifty times its diameter.

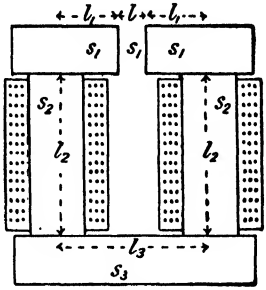


FIG. 282.

(iii) *Core of Electro-magnet.*—In a complicated case such as the core of an electro-magnet (Fig. 282), an approximate value

of the magnetic resistance may be obtained from the dimensions of the circuit. Thus—

for the air gap, magnetic resistance $=\dfrac{l}{s_1}$,

for the pole pieces ,, ,, $=\dfrac{2l_1}{\mu_1 s_1}$,

for the cores ,, ,, $=\dfrac{2l_2}{\mu_2 s_2}$,

and for the yoke ,, ,, $=\dfrac{l_3}{\mu_3 s_3}$.

The total magnetic resistance is then—

$$\frac{l}{s_1}+\frac{2l_1}{\mu_1 s_1}+\frac{2l_2}{\mu_2 s_2}+\frac{l_3}{\mu_3 s_3},$$

and if there are nI ampere turns—

$$\frac{4\pi n\mathrm{I}}{10}=\mathrm{N}\left(\frac{l}{s_1}+\frac{2l_1}{\mu_1 s_1}+\frac{2l_2}{\mu_2 s_2}+\frac{l_3}{\mu_3 s_3}\right).$$

The value of B in the air-gap is N/s_1, and since the permeability is here unity, this is also the strength of field in the gap. The above calculation is only an approximation, since the circuit is very imperfect from the magnetic point of view, but it becomes more reliable as the air-gap is reduced, since the magnetic induction in that case will be more confined to the iron, less straying into the surrounding space.

Bar and Yoke Tests.—The objections to the magnetometer and the ring methods of measurement of permeability, consist in the necessity for drawing the material into the form of a wire in the former case and welding it into a ring in the latter. Both processes produce physical change in the material, and hence the desirability of employing some method in which these processes are unnecessary. Hopkinson [1] used a straight bar of the material and completed the magnetic circuit by means of a heavy soft-iron yoke, so that there are no free poles (Fig. 283). The magnetising coil is wound on the rod, the number of ampere-turns being known. In the earlier experiments, the experimental rod was constructed in two parts so that the secondary coil S, which is in series with the ballistic

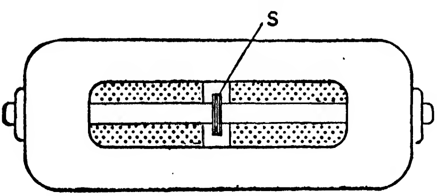


FIG. 283.

[1] J. Hopkinson, *Phil. Trans.*, 176, p. 455. 1885.

galvanometer, might be jerked out of the field on separating the parts of the rod. The joint, where the ends of the rod are in contact, introduces an unknown magnetic resistance, so in later experiments the rod was made in one piece and the ballistic throw for a reversal of the magnetising current observed.

Then, if l_1, s_1 and μ_1 are the length, area and permeability of the rod, and l_2, s_2 and μ_2 the values for the yoke—

$$\text{magnetic resistance of circuit} = \frac{l_1}{\mu_1 s_1} + \frac{l_2}{\mu_2 s_2},$$

and,
$$\frac{4\pi n\mathrm{I}}{10} = \mathrm{N}\left(\frac{l_1}{\mu_1 s_1} + \frac{l_2}{\mu_2 s_2}\right),$$

where $n\mathrm{I}$ is the number of ampere-turns in the magnetising coil.

Then if H is the magnetising field, and B the induction in the rod—

$$\mathrm{N} = \mathrm{B}s = \mu_1 \mathrm{H} s_1 ;$$

$$\therefore \mathrm{H} = \frac{4\pi n\mathrm{I}}{10\left(l_1 + \frac{l_2 s_1 \mu_1}{\mu_2 s_2}\right)}.$$

If the rod alone formed a complete magnetic circuit so that there were no free poles, as in the case of the ring, we should have had—

$$\mathrm{H} = \frac{4\pi n\mathrm{I}}{10 l_1},$$

and we see that the yoke is equivalent to an additional length $\frac{l_2 s_1 \mu_1}{\mu_2 s_2}$ of rod. This is made as small as possible by making s_2 and μ_2 large, the yoke being of high permeability soft iron, and as massive as possible.

Double Bar and Yoke.—The method has been modified by Sir J. A. Ewing [1] by using a double bar and yoke in such a way that the error due to the yoke is eliminated.

The above equation for H may be written—

$$\mathrm{H} = \frac{4\pi n\mathrm{I}}{10 l_1} - \frac{l_2 s_1 \mu_1 \mathrm{H}}{s_2 \mu_2 l_1}.$$

But $\mu_1 \mathrm{H} = \mathrm{B}$, the induction in the rod ;

$$\therefore \mathrm{H} = \frac{4\pi n\mathrm{I}}{10 l_1} - \mathrm{B}\left(\frac{l_2 s_1}{s_2 \mu_2}\right)\frac{1}{l_1}.$$

Two rods RR are employed, which are united at their ends by two massive soft-iron yokes YY (Fig. 284). Then, for any given value of B, the quantity $\mathrm{B}\left(\frac{l_2 s_1}{s_2 \mu_2}\right)$ is constant for the given yokes,

[1] Ewing, "Magnetic Induction in Iron and other Metals."

and writing ϵ in place of it, we have, taking l as the length of the rod—

$$H=\frac{4\pi nI}{10l}-\frac{\epsilon}{l}.$$

Now $\frac{4\pi nI}{10l}$ is the magnetising field H' due to the coil when there

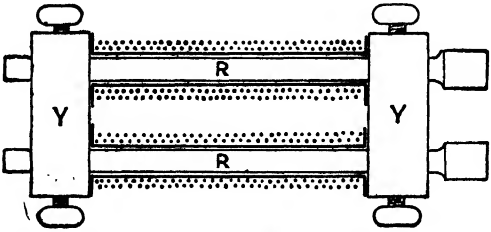


FIG. 284.

is no correction to be applied on account of the magnetic resistance of the yokes.

$$\therefore\ H=H'-\frac{\epsilon}{l}.$$

With half the length of rod, the equation would have been—

$$H=H''-\frac{2\epsilon}{l}$$

$$\therefore\ H''-H'=\frac{\epsilon}{l}$$

$$=H'-H.$$

The measurement of the induction B is therefore made twice over, the length of rod in the second case being half the length in the first, with the same number of turns per centimetre of the magnetising coil in the two cases. This is attained by having the magnetising coils wound on two pairs of bobbins, one twice the length of the other, the first having 100 turns on 12·56, *i.e.* 4π cm., and the other 50 turns on 6·28, *i.e.* 2π cm., so that in each case $\frac{4\pi nI}{10l}=10\times I$.

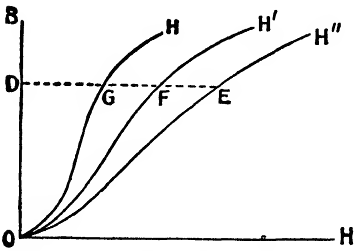


FIG. 285.

A curve of B and H for each arrangement is obtained by the ballistic method in the ordinary way, and the curves for H′ and

H″ plotted as in Fig. 285. Then for each value of B, such as OD, we have seen that—

$$H''-H'=H'-H.$$

Hence if FG be made equal to EF, *i.e.* to H″−H′, the point G is situated on the true or corrected B—H curve.

The two bars must be of the same material, and the B—H curve being found, one of the bars may be compared with a bar of any other material, having the same dimensions, by the rapid method to be next described. The experiment with the double bar and yoke is carried out in a similar manner to that with the ring (p. 276), a set of observations being made with each pair of magnetising coils. The secondaries of these and of the standardising coil are permanently in series, so that the resistance of the secondary circuit is not changed during the experiment.

Permeability Bridge.[1]—The rod A, standardised by the double bar and yoke method, and B the rod to be tested, are placed between massive yokes CD as in the double bar and yoke test, the direction of magnetisation produced by the magnetising coils being shown by arrows in the plan in Fig. 286. The distance between the yokes is 12·56 cm., and the number of turns upon A, 100, so that the magnetising field in A is 10×I. The number of turns upon B can be varied by means of a switch or plugs, until the induction is the same in the two rods, the ratio of the number of magnetising turns upon A and B when this balance is obtained being the ratio of the two magnetising fields required to produce that particular value of B in the two rods. If M (Fig. 287) is a point on the B—H curve found by the last experiment (p. 302) for the rod A, and the ratio LN : LM be that of the magnetising fields in B and A for this value of the induction, the point N on the B—H curve for the specimen is found. The whole curve ONH_1 may be found from the standard curve OMH in the same manner by taking the ratio of the magnetising turns for various values of the induction B.

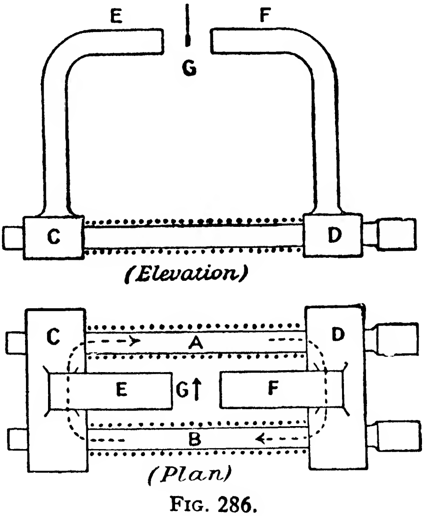


FIG. 286.

[1] J. A. Ewing, *The Electrician*, 37, p. 41. 1896.

The test for equality of the values of B in the two rods is made by observing that the suspended magnet G (Fig. 286) remains stationary when the magnetising current, which passes through the two magnetising coils in series, is reversed. This shows that there is no magnetic flux passing from C to D through the soft-iron horns E and F. Hence the flux entering the yoke C from the specimen B is equal to that leaving C by the specimen A, and the same for the yoke D, and the magnetic induction in A is therefore equal to that in B.

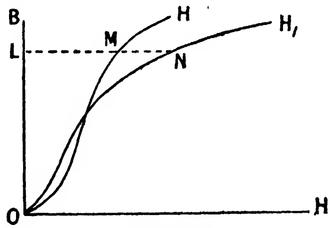


FIG. 287.

Magnetic Analysis.—Steel is a mixture of iron and carbon of varying complexity, dependent upon temperature. One constituent is iron carbide (Fe_3C), which has a magnetic critical temperature of 213° C. It has been shown by S. W. J. Smith [1] that on measuring the magnetisation of steel during heating or cooling, a small change in magnetisation occurs at 213° C. The amount of the change can be used to measure the proportion of carbide in the steel. By using an appropriate strength of field the change can be made very abrupt. On examining the critical transition at 700°–800° C., Smith and Guild [2] found a combination of iron and carbon (eutectoid) to have a critical temperature different from the rest of the iron, and at such a sharply marked temperature (730° C.) that this might be used as a calibration temperature in thermometry. The change on cooling, when the eutectoid became magnetic, was not so sharply marked as that on heating, and occurred at a temperature below 730° C. Smith has pointed out the value of this magnetic analysis in determining the composition of steels. It is very delicate, and, unlike chemical analysis, it does not destroy the specimen.

[1] S. W. J. Smith, *Proc. Phys. Soc.*, xxv. p. 77. 1912.
[2] Smith and Guild, *Phil. Trans.*, ccxv. (A), p. 177. 1915.

CHAPTER X

VARYING CURRENTS

THE laws which govern the flow of steady currents in conductors are inadequate to determine the value of the current when changing. From Ohm's law we can define the resistance of a conductor to steady current, and then calculate the electromotive force required to produce any current in the conductor. When, however, the current varies, the magnetic flux linked with the circuit varies, and we have seen that this variation in the magnetic flux linked with the circuit means another electromotive force acting in it. Again, there may be accumulation of electric charge at some part in the circuit, and this also implies a changing electromotive force in the circuit. For simplicity, we shall in the first place neglect this last effect, which is only of importance within certain limiting values of the capacity in the circuit, and confine our attention to the effect of the changing magnetic flux. The effect of capacity will afterwards be treated, and finally the circuit will be dealt with, in which the effects are all taken into account.

Inductance.—When the magnetic flux linked with a circuit is changing, an electromotive force acts round the circuit, whose value is the rate of change of the flux (p. 251). Thus—

$$e=-\frac{dN}{dt}.$$

Again, for current i in the circuit there is always a magnetic flux linked with the circuit; let this flux be li. Then—

$$N=li \quad \ldots \ldots \ldots \ldots \quad \text{(i)}$$

and,
$$e=-\frac{d(li)}{dt}=-l\,.\frac{di}{dt} \quad \ldots \ldots \quad \text{(ii)}$$

provided that l is constant, which is true so long as there is no material of variable permeability in the neighbourhood of the circuit.

When the current i is increasing, we may, by the laws on pages 251 and 252, show that the induced electromotive force e is in the opposite direction to the current; consequently work is being

performed at the rate ei ergs per second, the energy appearing somewhere in the circuit (p. 55). That is—

$$\text{rate of working} = ei$$
$$= -li \cdot \frac{di}{dt} \text{ ergs per second}$$

and for the total work done in opposing this E.M.F. while the current i_0 is being established—

$$-\int_0^t li\frac{di}{dt}dt = -l\int_0^{i_0} idi = -\tfrac{1}{2}li_0^2 \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

i_0 being the final steady value of the current.

The quantity l which appears in these three equations is called the *Coefficient of Self-Induction* or the *Self-Inductance* of the circuit; it may be defined from (i) as *the flux linked with the circuit when unit current flows in it;* from (ii) as *the E.M.F. round the circuit due to unit rate of change of current in it;* and from (iii) as *twice the work done in establishing the magnetic flux associated with unit current in the circuit.*

The three values of l are constant, and are, moreover, identical so long as the medium comprising the magnetic circuit linked with the current has constant magnetic permeability, but when μ is variable the three values are neither identical nor constant, and the inductance of the circuit may be defined from either of the equations, the question of the most convenient definition for any particular problem being decided by experience. Thus, if OEA (Fig. 288) be the curve connecting N the total magnetic flux and the current i in the circuit considered, it is of the same form as the B—H curve, the scale only being changed; for N$=$Bs, when B is constant across any section of the magnetic circuit, or N$=\int$Bds in any case; and the magnetising field is everywhere proportional to the current. Then from equation (i)—

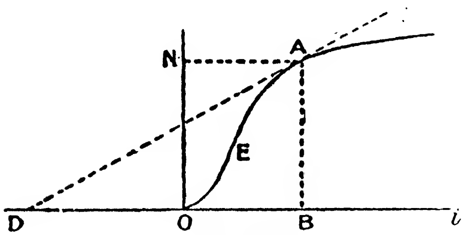


FIG. 288.

$$l = \frac{\text{N}}{i} = \frac{\text{AB}}{\text{OB}}.$$

Again,
$$e = -\frac{d\text{N}}{di} \cdot \frac{di}{dt},$$

and $\frac{d\text{N}}{di}$ is equal to $\frac{\text{AB}}{\text{DB}}$, where AD is the tangent to the curve at

the point A ; hence—

$$e=-\frac{\mathrm{AB}}{\mathrm{DB}}\cdot\frac{di}{dt}.$$

But from equation (ii)—

$$e=-l\frac{di}{dt},$$

$$\therefore\ l=\frac{\mathrm{AB}}{\mathrm{DB}}.$$

The work done in establishing the current is—

$$\int_0^t eidt=-\int_0^t i\frac{d\mathrm{N}}{dt}dt=-\int_0^{\mathrm{N}} id\mathrm{N}$$

$$=\text{area OEANO}.$$

But from equation (iii)—

$$\text{work}=-\tfrac{1}{2}l\,.\,\mathrm{OB}^2,$$

$$\therefore\ \tfrac{1}{2}l\,.\,\mathrm{OB}^2=\text{area OEANO}.$$

$$\therefore\ l=\frac{2\text{ area OEANO}}{\mathrm{OB}^2}.$$

It will easily be seen that if OEA becomes a straight line, in which case the permeability is constant, all these quantities are equal and are moreover constant.[1]

In practice the meaning of the term self-inductance must vary with the conditions in which it is employed ; with iron in the circuit it varies from value to value of the current, but without iron the term has a definite meaning, and for a given conducting circuit it is constant so long as the current does not change so rapidly that the distribution of current in the conductor itself differs from that for a steady current.

Growth of Current.—While the current in a circuit is growing, the electromotive force $-l\frac{di}{dt}$ is acting, and therefore the resultant electromotive force overcoming the resistance of the circuit is $e-l\frac{di}{dt}$, where e is the applied electromotive force due to outside sources. Hence our equation from which to obtain the current, changes from $e=ri$ for steady current, to $e-l\frac{di}{dt}=ri$ for varying current,

$$\therefore\ l\frac{di}{dt}+ri=e.$$

[1] W. E. Sumpner, *Phil. Mag.* (Ser. 5), 25, p. 453. 1888.

This is the general equation of electromotive forces for a circuit having inductance and resistance only. In order to integrate it let it be written in the form—

$$\frac{di}{\frac{e-ri}{l}}=dt,$$

or, assuming l to be constant—

$$-\frac{l}{r}\cdot\frac{d\left(\frac{e-ri}{l}\right)}{\frac{e-ri}{l}}=dt.$$

Integrating both sides, we have—

$$\frac{l}{r}\log_\epsilon\left(\frac{e-ri}{l}\right)=-t+k.$$

k is the constant of integration, to be determined from the conditions of the problem; for the differential equation represents the manner in which the quantities vary with respect to each other, and its integral is the total change in a given time, but will not give us the value of the current reached at the end of this time, unless we know the value at the beginning of the time during which the change has taken place. Let the current at the moment of applying the external E.M.F. be zero, *i.e.* let $i=0$ when $t=0$. Then—

$$\frac{l}{r}\log_\epsilon.\frac{e}{l}=k,$$

and substituting this value for k we have—

$$\frac{l}{r}\log_\epsilon\left(\frac{e-ri}{l}\right)-\frac{l}{r}\log_\epsilon\frac{e}{l}=-t$$

or,
$$\log_\epsilon\frac{e-ri}{e}=-\frac{r}{l}t.$$

Writing this in its exponential form we have—

$$\frac{e-ri}{e}=\epsilon^{-\frac{r}{l}t}$$

or,
$$i=\frac{e}{r}(1-\epsilon^{-\frac{r}{l}t})$$

$\frac{e}{r}$ is the final steady value of the current, and writing i_0 for this we have—

$$i=i_0(1-\epsilon^{-\frac{r}{l}t}),$$

an equation which shows us how the current grows. The mode of growth is shown in Fig. 289, in which the values of i and t are plotted. Strictly speaking the current never reaches its steady value i_0 but continually approaches it.

Thus for i to equal i_0—

$$\epsilon^{-\frac{rt}{l}}=0,$$
$$\therefore t=\infty.$$

The rate of growth of the current may easily be found ; it is

$$\frac{di}{dt}=\frac{r}{l}i_0\epsilon^{-\frac{rt}{l}}=\frac{r}{l}(i_0-i).$$

This gets less as i approaches its final value i_0, but for any value of i it is proportional to $\frac{r}{l}$. The rate at which a current approaches its final value therefore depends upon the ratio $\frac{r}{l}$ and

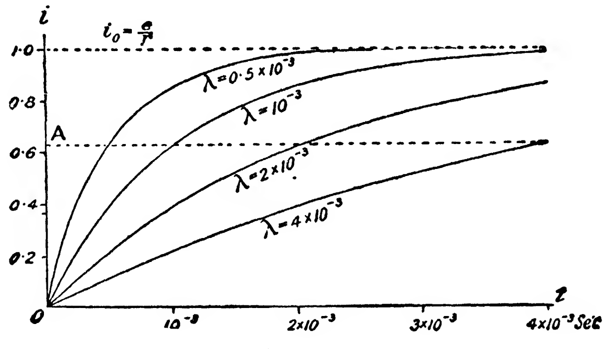


FIG. 289.

not upon the separate values of r and l. Unless the circuit has a great many turns, or the circuit includes iron, l is usually very small compared with r, so that the current approximates to its final value in a very small fraction of a second.

The ratio $\frac{l}{r}$ is called the *Time Constant*, λ, of the circuit, and the equation for the current may be written—

$$i=i_0(1-\epsilon^{-\frac{t}{\lambda}}).$$

After time λ,
$$i=i_0\left(\frac{\epsilon-1}{\epsilon}\right)$$
$$=i_0\cdot\frac{1·718}{2·718}=0·632i_0.$$

Thus the time constant is the time in which a current reaches $\frac{1{\cdot}718}{2{\cdot}718}$ or roughly two-thirds of its final value. The four curves in Fig. 289 are drawn for values of the time constant equal to 4×10^{-3}, 2×10^{-3}, 10^{-3}, and $0{\cdot}5\times10^{-3}$, and the time taken for each current to reach the line A, where $OA=0{\cdot}632i_0$, is the time constant.

The variation in rate of growth of the current cannot be observed with an ordinary galvanometer, but by using one having a very high frequency for its moving part, as for example the vibration galvanometer (p. 387) the difference in rates of growth of current in an electromagnet and in an equal resistance having small inductance may easily be demonstrated.

Decay of Current.—If, when the steady value i_0 of the current has been reached, the electromotive force is suddenly reduced to zero, the variation of the current may be found, for e in our E.M.F. equation is then zero.

$$l\frac{di}{dt}+ri=0.$$

Transforming the equation as before, we get—

$$-\frac{l}{r}\cdot\frac{di}{i}=dt$$

and integrating, $\frac{l}{r}\cdot\log_\epsilon i=-t+k.$

When $t=0$, $i=i_0$.

$$\therefore \frac{l}{r}\log_\epsilon i_0=k,$$

and substituting this value for k, we have—

$$\log_\epsilon\frac{i}{i_0}=-\frac{r}{l}t,$$

$$i=i_0\epsilon^{-\frac{r}{l}t}$$

or, $$i=i_0\epsilon^{-\frac{t}{\lambda}}.$$

Thus the greater the value of λ, the more slowly will the current die away. Fig. 290 illustrates the decay of the current for $\lambda=4\times10^{-3}$, 2×10^{-3}, 10^{-3}, and $0{\cdot}5\times10^{-3}$.

It will be noticed that the growing and decaying currents are complementary, for if the currents $i_0(1-\epsilon^{-\frac{t}{\lambda}})$ and $i_0\epsilon^{-\frac{t}{\lambda}}$ be added together, the sum is i_0. Hence the curves in Fig. 290 are those of Fig. 289 inverted.

The variation in rate of decay of the current in different circuits may be demonstrated by means of the vibration galvano-

meter in a manner similar to that described for the growing current (p. 310), but in this case the circuit must be closed as the battery is cut out ; it must not merely be broken, as this would introduce an infinite resistance. The result of this increase in resistance is that the time constant of the circuit is reduced and the current drops at an enormous rate. The electromotive force due to inductance, being proportional to the rate of change of current, is then enormous, and is sufficient to cause a spark or even an arc at the break in the circuit.

Inductance in electrical problems plays a similar part to mass or inertia in mechanical problems ; its effect is to retard the

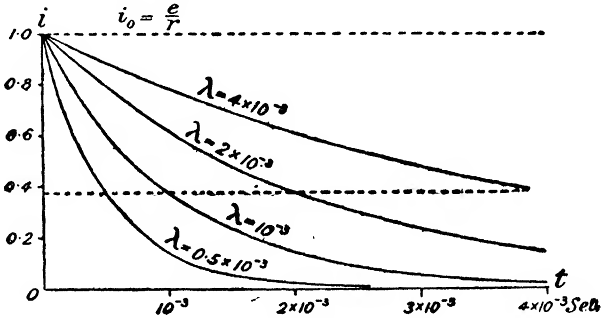


FIG. 290.

growth of the current or the motion, and similarly neither the current nor the motion can be stopped instantaneously. The energy due to the magnetic field linked with the current is $\frac{1}{2}li^2$ ergs ; that due to the inertia of a moving mass is $\frac{1}{2}mv^2$ ergs.

Inductance of Solenoid.—In certain simple cases, the self-inductance of a circuit may be calculated from the definition $e = -l \cdot \frac{di}{dt}$, provided of course that the permeability is constant. In the case of a solenoid having an air core, if b is its length, n the total number of turns, and a its area of cross-section—

$$\text{Field inside solenoid} = \frac{4\pi ni}{b}, \qquad \text{(p. 232)}$$

$$\text{magnetic flux} = \frac{4\pi nai}{b}.$$

If the solenoid is straight b must be great.

When i varies, any change in the flux cuts the circuit n times,

$$\therefore e = -n\frac{d}{dt}\left(\frac{4\pi nai}{b}\right)$$

$$= -\frac{4\pi n^2 a}{b} \cdot \frac{di}{dt}.$$

By comparison with the above equation for l we see that—

$$l=\frac{4\pi n^2 a}{b}.$$

Coaxial Cylinders.—When the circuit consists of two coaxial cylinders of radii a and b, one being the return circuit for the other, the magnetic field is confined to the space between them. For, a circular path taken externally round them both encloses equal and opposite currents (Fig. 291) and, therefore, the line integral of the magnetic field round the path is zero. Again, within the inner cylindrical current, the magnetic field is zero, for a closed path in this region does not enclose any current. It follows that the resultant magnetic field is confined to the space between the cylinders and is that due to the current i in the inner cylinder. At external points, the magnetic fields due to the two cylinders are equal and opposite. This proves incidentally that the field due to a cylindrical current is the same at external points, as though the current flowed along the axis, for the inner cylinder may be reduced to as small dimensions as we please.

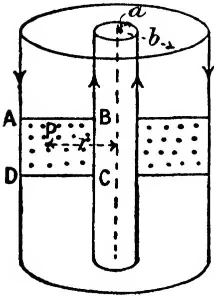


FIG. 291.

The value of H at the point P is therefore $\frac{2i}{r}$, and the flux through the area ABCD, where AD has unit length, is—

$$\int_a^b \frac{2i}{r}dr=2i\Big[\log_\epsilon r\Big]_a^b.$$

$$N=2i\log_\epsilon\frac{b}{a}.$$

Now when i varies—

$$e=-\frac{dN}{dt}$$

$$=-2\log_\epsilon\frac{b}{a}\cdot\frac{di}{dt}.$$

But, $$e=-l\frac{di}{dt}.$$

$$\therefore\ l=2\log_\epsilon\frac{b}{a}$$

per unit length of the coaxial cylinders.

Practical Unit (the Henry).—In the foregoing equations, all the quantities have been given in absolute C.G.S. units, but in

practical work it is desirable to employ a unit in conformity with the system—volt, ampere, ohm etc. The unit chosen is called the *Henry*, and is the inductance of a circuit in which a rate of change of current of one ampere per second produces an electromotive force of one volt. The relation of the henry to the absolute unit of inductance may be found in a manner analogous to that employed in the case of the ohm (p. 58). The volt is equal to 10^8 absolute units of electromotive force and the ampere to 10^{-1} unit of current, whereas the second is the unit of time on both systems, and thus, since—

$$e=l\frac{di}{dt};\quad L=\frac{\text{E volts}}{\frac{dI}{dt}\text{ amperes per sec.}}\text{ henries,}$$

or,
$$1\text{ henry}=\frac{1\text{ volt}}{1\text{ ampere per sec.}}$$
$$=\frac{10^8\text{ absolute units of E.M.F.}}{10^{-1}\text{ absolute unit of current per sec.}}$$
$$=10^9\text{ absolute units of inductance.}$$

Thus the inductance of a solenoid—

$$=\frac{4\pi n^2 a}{10^9 b}\text{ henries}\qquad\text{(p. 312)}$$

and that of the coaxial cylinders on p. 312 is $2\times10^{-9}\log_e\frac{b}{a}$ henry.

In future we shall use the letter l to represent inductance in absolute units, and L that in henries.

A convenient form of variable inductance made by Messrs. Nalder Bros. & Co., Ltd., is shown in Fig. 292. The two coils are in series, and one of them can be rotated so as to vary the resultant magnetic flux due to the two. The scale on the instrument is calibrated in millihenries.

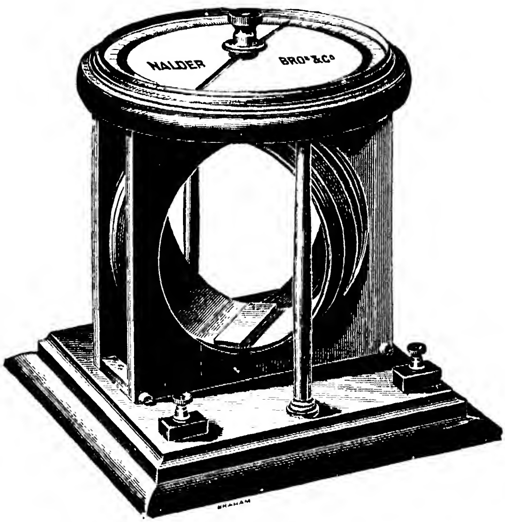


FIG. 292.

Charge and Discharge of Condenser.—On applying an electro-

motive force e, due to some external source, to a circuit consisting of a capacity in series with a resistance, a current will flow for a time, but electric charge is all the time accumulating upon the plates of the condenser, and for charge q, the difference of potential between the plates is $\frac{q}{c}$, where c is the capacity of the condenser. This is directed one particular way round the circuit (Fig. 293), and is equivalent to an electromotive force. A current in one direction will increase this, giving energy to the condenser, and for a current in the other direction, the energy of the charge on the condenser will be used in driving the current. Hence the condenser produces an electromotive force in the circuit, and the equation of electromotive forces becomes—

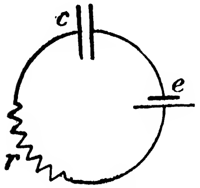


FIG. 293.

$$ri+\frac{q}{c}=e.$$

Now the current in every part of the circuit being the same, it is equal to the rate at which charge accumulates in the condenser.

$$i=\frac{dq}{dt}, \text{ or, } q=\int idt,$$

$$\therefore\ r\frac{dq}{dt}+\frac{q}{c}=e.$$

This equation may be solved in a similar manner to that on p. 308.

$$\text{Thus, } -cr\,.\frac{d\left(\frac{e-\frac{q}{c}}{r}\right)}{\left(\frac{e-\frac{q}{c}}{r}\right)}=dt$$

$$cr\log_\epsilon\left(\frac{e-\frac{q}{c}}{r}\right)=-t+k.$$

If $q=0$, when $t=0$—

$$cr\log_\epsilon\frac{e}{r}=k.$$

$$\therefore\ \log_\epsilon\frac{e-\frac{q}{c}}{e}=-\frac{t}{cr}$$

$$q=ec(1-\epsilon^{-\frac{t}{cr}})$$

ec is the final steady charge in the condenser ; calling this q_0 we have—

$$q=q_0(1-\epsilon^{-\frac{t}{cr}}).$$

The time constant λ in this case is cr,

$$\therefore\ q=q_0(1-\epsilon^{-\frac{t}{\lambda}}).$$

If now the external electromotive force be reduced to zero, the E.M.F. equation becomes—

$$r\frac{dq}{dt}+\frac{q}{c}=0,$$

whence,

$$\frac{dq}{q}=-\frac{dt}{cr}$$

$$\log_\epsilon q=-\frac{t}{cr}+k.$$

If $q=q_0$, when $t=0$—

$$k=\log_\epsilon q_0$$

or,

$$\log_\epsilon \frac{q}{q_0}=-\frac{t}{cr},$$

$$\therefore\ q=q_0\epsilon^{-\frac{t}{cr}}=q_0\epsilon^{-\frac{t}{\lambda}}.$$

The curves for q and t for charge and discharge are drawn in Fig. 294.

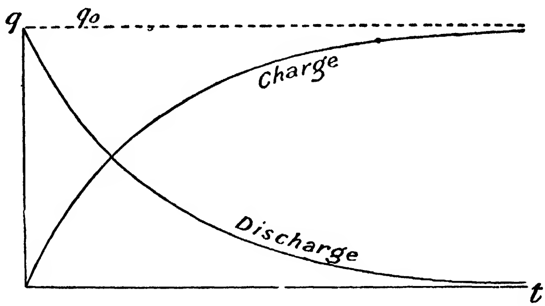


FIG. 294.

The equations for the current may be obtained from those for the charge, remembering that $i=\frac{dq}{dt}$.

Thus, for the charge—

$$i=q_0\frac{d}{dt}(1-\epsilon^{-\frac{t}{cr}})=\frac{q_0}{cr}\epsilon^{-\frac{t}{cr}}$$

and for the discharge—

$$i=q_0\frac{d}{dt}(\epsilon^{-\frac{t}{cr}})=-\frac{q_0}{cr}\epsilon^{-\frac{t}{cr}}.$$

Since $$e=\frac{q_0}{c},$$

$$\frac{q_0}{cr}=\frac{e}{r}=i_0,$$

i_0 being the value of the current at the beginning of both charge and discharge. It may be noticed that in both cases the current starts with its greatest value and falls off exponentially; the discharging current is negative, that is, it is in the reverse direction to the charging current (Fig. 295).

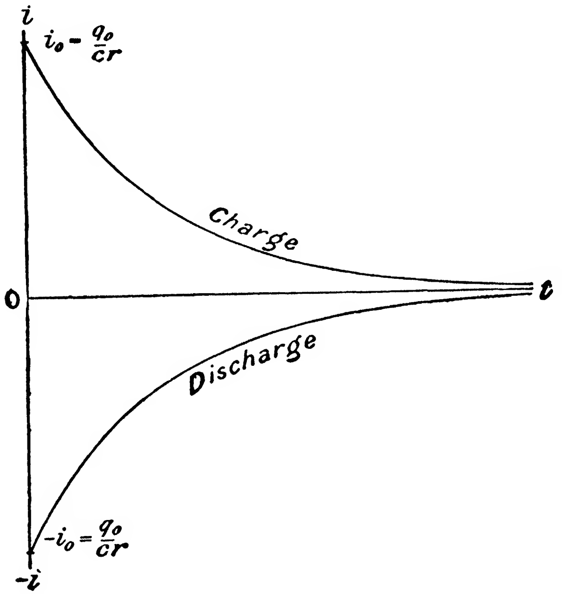


FIG. 295.

Measurement of High Resistance by Leakage.—From the equation for the discharge, we see that q falls to about one-third of its value in time $\lambda=cr$; for, $\frac{q}{q_0}=\epsilon^{-\frac{t}{\lambda}}=\frac{1}{\epsilon}$, when $t=\lambda$. If the time t is to be measured with reasonable accuracy, it must be over a minute, say 100 seconds. Therefore cr must be at least 100. Now the condensers of convenient size, found in every laboratory, have a capacity of the order of a micro-farad, that is 10^{-6} farad, or 10^{-15} absolute units, where the practical unit of capacity, the *Farad*, is the condenser which one coulomb will charge to a

potential difference of one volt. Hence for cr to have the value 100, when c is 10^{-15}, r must have a value of 10^{17} absolute units or 10^8 ohms. Hence a capacity of 1 micro-farad discharging through a resistance of 10^8 ohms or 100 megohms (1 megohm $=10^6$ ohms) will lose about two-thirds of its charge in 100 seconds.

This gives rise to a convenient practical method of measuring resistances of the order of 20 megohms and upwards; for the condenser is charged and then allowed to discharge through the resistance for a known time t. From the relation—

$$\log_\epsilon \frac{q}{q_0} = -\frac{t}{cr},$$

or,

$$r = \frac{t}{c \log_\epsilon \dfrac{q_0}{q}},$$

c being known and $\dfrac{q_0}{q}$ being observed, r can be calculated. When the rate of leakage is small, $\dfrac{q_0}{q}$ may be measured by means of the quadrant electrometer, the deflection being read at known intervals.

$$\frac{\theta_1}{\theta_2} = \frac{e_0}{e} = \frac{q_0}{q}.$$

Or the condenser may be charged and instantly discharged through the ballistic galvanometer. The throw θ_1 is then proportional to q_0. It is again charged and allowed to leak for t seconds through the resistance and then discharged through the galvanometer. The throw θ_2 is then proportional to the charge q remaining after t seconds, so that—

$$r = \frac{t}{c \log_\epsilon \dfrac{\theta_1}{\theta_2}}.$$

It is advisable to obtain a number of readings of t and θ by repeating the above process and plotting a curve of t and $\log_\epsilon \dfrac{\theta_1}{\theta_2}$. A straight line lying evenly amongst these points may be drawn and from it a mean value of $\dfrac{t}{\log_\epsilon \dfrac{\theta_1}{\theta_2}}$ obtained, from which r may be calculated. It should be remembered that the dielectric of the condenser may not be a perfect insulator. There will then be a leak, even when there is no external conductor. Hence a measurement should be made when the terminals of the con-

denser are unconnected and the resistance r_1 of the condenser leak be found. Then with the high resistance across the terminals the combined resistance r_2 should be found. Since the high resistance and the condenser are in parallel

$$\frac{1}{r}+\frac{1}{r_1}=\frac{1}{r_2}$$

from which r can be calculated.

Mutual Inductance.—We have already seen that a variation of the current in a circuit is accompanied by an electromotive force in any neighbouring circuit (p. 250). Thus if the current in the circuit A (Fig. 239) varies, there will be an electromotive force in the circuit B, equal to $-m\frac{di}{dt}$, due to this variation of the current in A. m is called the *coefficient of mutual induction* or the *mutual inductance* of the two circuits. The defining of mutual inductance is subject to all the difficulties encountered in the case of self-inductance when the magnetic permeability is variable (p. 306). It may be defined as above, or as the magnetic flux linked with the secondary circuit B, due to unit current in the primary A. Thus—

$$e=-\frac{dN}{dt}=-\frac{d(mi)}{dt}=-m\frac{di}{dt},$$

when m is constant.

m may also be defined as the mutual potential energy of the two circuits when unit current is flowing in each, and this again leads to the same value of m when the permeability is constant.

Let i_2 be the current in the secondary circuit, and let it be situated in a magnetic field whose value at the point P, Fig. 296 (i), is H, θ being the angle between the field and the circuit at P. Then the force per unit length of the circuit is $i_2\mathrm{H}\sin\theta$ (see p. 239), and for the small length l of the circuit, is $i_2\mathrm{H}l\sin\theta$, and is at right angles to l and H. Let the element l be displaced in the direction of the force by an amount δx, then work done $=i_2\mathrm{H}\sin\theta.l.\delta x$.

But H sin θ is the component of H normal to the area $l\,.\,\delta x$ swept out by the element (Fig. 296 (ii)), and therefore $\mathrm{H}\sin\theta\, l\,.\,\delta x$ is the amount of magnetic flux $d\mathrm{N}$ added to or withdrawn from the total amount linked with the circuit. Hence—

$$\text{work done}=i_2 d\mathrm{N},$$

and $\int i_2 d\mathrm{N}$ is the total work done when the magnetic flux N, linked with the circuit, changes by the amount $\int d\mathrm{N}$, i_2 being constant.

It is immaterial whether the circuit changes in size, or whether the change in N is due to the alteration in the distribution of the

flux, for the work done at each element of the circuit is proportional to the value of H at the point, and to the relative motion of the circuit and the flux; in fact the problem is similar to that of finding the work done in the case of the change of volume of a gas, which is $\int p\,dv$, where p is the pressure at the boundary and dv a small change in volume. The same limitations as regard reversibility apply in the two cases. If the permeability is not unity H must be replaced by B in the above.

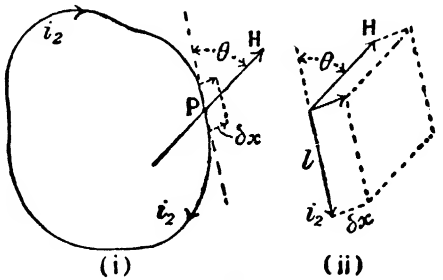


FIG. 296.

If now the flux N is due to another circuit carrying current i_1, the flux due to it and linked with both circuits is $m_1 i_1$, and therefore $m_1 i_1 i_2$ is the work done in linking the flux $m_1 i_1$ with, or withdrawing it from i_2. Similarly the current i_2 involves a flux $m_2 i_2$ linked with i_1, and to withdraw this flux from i_1 involves an amount of work $m_2 i_2 i_1$. These two amounts of work must be the same, for if the two circuits be separated to a great distance, the forces on the two at each instant during the act of separation must be equal and opposite.

$$\therefore\ m_1 i_1 i_2 = m_2 i_2 i_1$$

and,
$$m_1 = m_2,$$

so that there is only one value of the mutual inductance between the two circuits, and the flux linked with the second due to unit current in the first is equal to the flux linked with the first due to unit current in the second.

Calculation of Mutual Inductance.—In any case in which the flux linked with the secondary circuit due to current i in the primary circuit can be calculated, the mutual inductance may be deduced from the relation $e = -m\frac{di}{dt}$. Thus for the solenoid (Fig. 251) in which n turns of secondary are wound near the middle of a primary of n_1 turns per unit length—

$$H = 4\pi n_1 i, \text{ and, } N = 4\pi n_1 A i.$$

This is the flux linked with each turn of the secondary.

$$\therefore\ e = -n_2 \cdot \frac{d}{dt}(4\pi n_1 A i)$$

$$= -4\pi n_1 n_2 A \cdot \frac{di}{dt},$$

from which,
$$m = 4\pi n_1 n_2 A.$$

If the flux due to the current in the primary is not all linked with the secondary, as, for example, when the secondary turns are not all wound near the middle of the solenoid, the mutual inductance will be less than the above amount.

In the case of two identical circuits wound so that they practically coincide with each other everywhere, the mutual inductance would be equal to the self-inductance of either.

From the identity in form of the quantities self, and mutual, inductance they are measured in the same units. Thus the henry is the practical unit of mutual inductance, and is the mutual inductance of a pair of circuits when a rate of change of one ampere per second in one, causes an electromotive force of one volt in the other. We shall write "m" for mutual inductance measured in absolute units, and "M" for that measured in henries.

In comparing inductances experimentally, it is often convenient to have a variable standard of inductance, but in the case of self-inductance there is the difficulty that the low values cannot be obtained, since, however the positions of the two parts of the circuit (see Fig. 292) are varied, the self-inductance can never be reduced to zero. Mr. A. Campbell[1] has suggested instead, the employment of standards of mutual inductance, since this can be varied for two coils from zero, or even a negative value, up to a maximum, by altering the relative positions of the primary and secondary coils.

Current in Secondary.—On starting the current in the primary circuit, we have seen that there is a current in the secondary, which ceases when that in the primary has become steady. Further, on stopping the primary current we again get a transient current in the secondary. To find the value of the current in the secondary at any moment, we must write the electromotive force equations for the two circuits and then obtain a solution. Let i_1, l_1 and r_1, i_2, l_2 and r_2 be the currents, inductances and resistances of the two circuits, and m the mutual inductance; then for the primary—

$$l_1\frac{di_1}{dt}+m\frac{di_2}{dt}+r_1i_1=e.$$

and for the secondary—

$$l_2\frac{di_2}{dt}+m\frac{di_1}{dt}+r_2i_2=0.$$

To obtain the equations for the primary and secondary currents, these two simultaneous equations must be solved. The mathematics involved is beyond the scope of this work, but the

[1] A. Campbell, *Proc. Roy. Soc.*, Ser. A., **79**, 428. 1907.

currents may be plotted by the step by step method with the help of these equations.

Writing them in the form—

$$\frac{l_1}{r_1}\frac{di_1}{dt}+\frac{m}{r_1}\frac{di_2}{dt}=\frac{e}{r_1}-i_1=i_0-i_1,$$

and

$$\frac{l_2}{r_2}\frac{di_2}{dt}+\frac{m}{r_2}\frac{di_1}{dt}=-i_2,$$

where i_0 is the final steady value of the current in the primary, we can then solve the simultaneous equations for $\frac{di_1}{dt}$ and $\frac{di_2}{dt}$. This gives us—

$$\begin{cases}\dfrac{di_1}{dt}=\dfrac{(i_0-i_1)l_2r_1+i_2mr_2}{l_1l_2-m^2}\\[2ex] \dfrac{di_2}{dt}=-\dfrac{(i_0-i_1)mr_1+i_2l_1r_2}{l_1l_2-m^2},\end{cases}$$

or,

$$\begin{cases}di_1=\dfrac{(i_0-i_1)l_2r_1+i_2mr_2}{l_1l_2-m^2}dt\\[2ex] di_2=-\dfrac{(i_0-i_1)mr_1+i_2l_1r_2}{l_1l_2-m^2}dt.\end{cases}$$

If small intervals of time, dt, are taken, we can begin with any values of i_1 and i_2 we please, say $i_1=0$ and $i_2=0$, and find the values of di_1 and di_2 for the first interval. From these we know the values of i_1 and i_2 for the beginning of the second interval and can then calculate di_1 and di_2 for the second interval. This process may be repeated until i_1 has reached its steady value and i_2 has again become zero. The first two curves in Fig. 297 have been obtained in this way, taking $L_1=10$ henries, $L_2=1$ henry and $M=0{\cdot}8$ henry, $R_1=10$ ohms, $R_2=1$ ohm, and $E=10$ volts, in which case the equations are written—

$$\begin{cases}dI_1=\dfrac{(I_0-I_1)L_2R_1+I_2MR_2}{L_1L_2-M^2}dt,\\[2ex] dI_2=-\dfrac{(I_0-I_1)MR_1+I_2L_1R_2}{L_1L_2-M^2}dt\end{cases}$$

and $I_0=1$ ampere. It will be seen that after six seconds the steady state has been very nearly reached. The dotted curve gives the growth of the primary current when there is no secondary circuit. The second curves have been drawn for the falling

primary current. They may be obtained by putting $E=0$ in the differential equations, whence—

$$\begin{cases} dI_1=\dfrac{-I_1L_2R_1+I_2MR_2}{L_1L_2-M^2}dt \\ dI_2=\dfrac{+I_1MR_1-I_2L_1R_2}{L_1L_2-M^2}dt, \end{cases}$$

and taking the initial value of I_1 to be I_0, *i.e.* $\dfrac{E}{R_1}$.

Charge Flowing in Secondary Circuit.—The quantity of electricity that has been caused to circulate in the secondary circuit

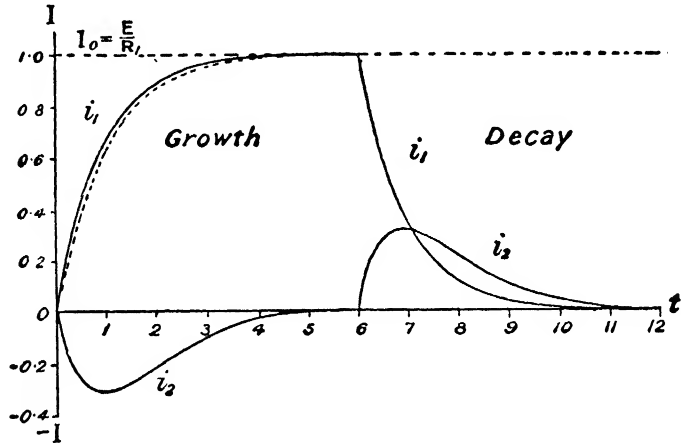


FIG. 297.

at either the starting or the stopping of the primary current, is $\int_0^\infty i_2dt$ and is the area included between the i_2 curve and the axis (Fig. 297). It may also be found from the equation—

$$\frac{l_2}{r_2}\cdot\frac{di_2}{dt}+\frac{m}{r_2}\cdot\frac{di_1}{dt}+i_2=0.$$

Integrating this with respect to time, from zero to infinity, we get—

$$\frac{l_2}{r_2}\int_0^\infty\frac{di_2}{dt}dt+\frac{m}{r_2}\int_0^\infty\frac{di_1}{dt}dt=-\int_0^\infty i_2dt.$$

Now at time 0, $i_2=0$, and, $i_1=0$,
and at time ∞, $i_2=0$, and, $i_1=i_0$.

Therefore the first term is zero at both limits, and—

$$\frac{m}{r_2}\int_0^{i_0}di_1=-\int_0^\infty i_2dt.$$

Now $\int_0^\infty i_2 dt$ is the quantity of electricity that has circulated in the secondary circuit, and therefore it is equal to $-\frac{mi_0}{r_2}$. When the current is stopped, the limits 0 and i_0 are reversed, and therefore the quantity is $\frac{mi_0}{r_2}$.

The Induction Coil.—A particular use of mutual inductance, of great practical importance, is made in the case of the induction coil, which is a piece of apparatus for producing small currents at very high electromotive force from comparatively large currents at low electromotive force. The primary coil PP (Fig. 298), consisting of a number of turns of thick wire, is wound upon an iron core D built up of a number of strands of soft-iron wire, while the secondary coil SS has a great number of turns of fine wire, and is wound upon the primary. On starting the current in the primary, the magnetic flux produced in the core cuts the secondary, producing a high electromotive force, and when the primary current is stopped, the flux again cuts the secondary but in the opposite direction, causing a reversed electromotive force. Many induction coils are provided with an automatic make and break, which consists of a spring having at its extremity a soft-iron armature A which is attracted towards the core when the primary current is made. This breaks the primary circuit at B, and on the core becoming demagnetised the spring recovers its original position and makes the circuit again by means of the contact B, and the process is then repeated. As considerable sparking occurs at B when the circuit is broken, the surfaces that come into contact are faced with platinum. This prevents undue sparking and wearing away.

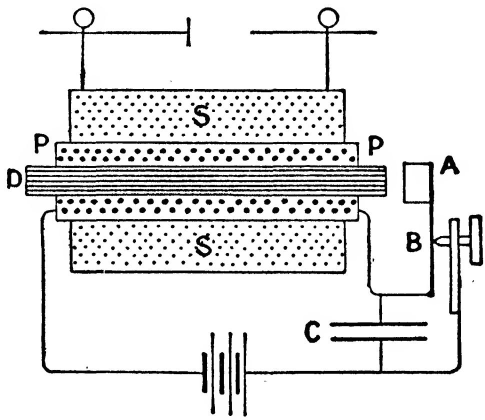


FIG. 298.

Although an electromotive force is produced in the secondary coil at both make and break of the primary circuit, the latter is by far the greater, since the primary current dies away much more rapidly than it grows. When the circuit is closed, its

resistance is small and its time constant $\left(\frac{l}{r}\right)$ great, but when the circuit is broken r is enormously increased and the time constant correspondingly reduced ; thus the rate of decay of the primary current is high. The magnetic flux in the core is therefore removed much more rapidly than it is produced, and the electromotive force in the secondary is much higher at the break of the primary circuit than at the make. On this account the electromotive force in the secondary circuit at the make is usually insufficient to produce a discharge through the gas surrounding the secondary terminals, and the secondary current is therefore unidirectional.

The manner in which the secondary current rises and falls on each break of the primary circuit may be seen from Fig. 297.

The efficiency of the coil is much increased if a condenser is placed in parallel with the contact breaker of the primary circuit, for the primary coil and the condenser comprise a circuit in which electrical oscillations occur, the condenser at the end of the first half oscillation being charged oppositely to its condition at the instant of break. The magnetic flux in the core due to the primary current is therefore reversed at each break, and the amount of charge caused to circulate in the secondary approaches double the value in this case, of that when no condenser is used, for without condenser the primary current merely drops to zero on account of the high resistance introduced at the break. The oscillations in the primary current will be rapidly damped, owing to the loss of energy due to heating produced by the currents in both primary and secondary, so that only the first discharge is of importance.

The effect of the condenser upon the secondary current may be found, on neglecting the effect of the resistance of the secondary at the beginning of the discharge. This is, to a first approximation, justified, for when the secondary current is varying rapidly, as at the beginning of the discharge, the predominant factor in determining its rate of growth is the large inductance of the secondary circuit.

The E.M.F. equations for the two circuits are therefore—

$$l_1\frac{di_1}{dt}+m\frac{di_2}{dt}+r_1i_1=0$$

and,

$$l_2\frac{di_2}{dt}+m\frac{di_1}{dt}=0.$$

Multiplying the first by l_2 and the second by m and subtracting, we have—

$$(l_1l_2-m^2)\frac{di_1}{dt}+l_2r_1i_1=0,$$

the solution to which is—

$$i_1 = i_0 \epsilon^{-\frac{l_2 r}{l_1 l_2 - m^2} t}$$ (see p. 308).

Writing a for $\frac{l_2 r_1}{l_1 l_2 - m^2}$, we have—

$$i_1 = i_0 \epsilon^{-at}$$

i_0 is here the current in the primary at the moment of breaking the circuit.

Integrating the E.M.F. equation for the secondary circuit, we have—

$$l_2 i_2 + m i_1 = k$$

where k is a constant.

$$\therefore\ l_2 i_2 + m i_0 \epsilon^{-at} = k.$$

If $i_2 = 0$ when $t = 0$, then—

$$k = m i_0$$

and,

$$i_2 = \frac{m}{l_2} i_0 (1 - \epsilon^{-at}).$$

As t increases to infinity this gets nearer and nearer to the value $\frac{m}{l_2} i_0$, but we must remember that this would only be true if the resistance of the secondary circuit were zero, which is far from being the case. From the start, the effect of the resistance is to decrease the current, and a short time after the electromotive force due to variation in the primary current has reached zero, the secondary current will also become zero. The value $\frac{m i_0}{l_2}$ is the limit which the secondary current cannot exceed, and would only reach if the resistance were zero.

When the resistance r_1 has been replaced by a condenser of capacity c, the maximum value of the secondary current may be obtained approximately by a method given by the late Lord Rayleigh,[1] which suggested the above. The electromotive force equation for the primary circuit being—

$$l_1 \frac{di_1}{dt} + m \frac{di_2}{dt} + \frac{q}{c} = 0,$$

or,

$$l_1 \frac{d^2 q_1}{dt^2} + m \frac{d^2 q_2}{dt^2} + \frac{q}{c} = 0,$$

and for the secondary,

$$l_2 \frac{di_2}{dt} + m \frac{di_1}{dt} = 0,$$

or,

$$l_2 \frac{d^2 q_2}{dt^2} + m \frac{d^2 q_1}{dt^2} = 0.$$

[1] Hon. J. W. Strutt. *Phil. Mag.* (Ser. 4), 39, p. 428. 1870.

Multiply the first by l_2 and the second by m, and subtract, and we get—

$$(l_1l_2-m^2)\frac{d^2q_1}{dt^2}+\frac{l_2}{c}q=0.$$

This is of the type $\frac{d^2x}{dt^2}+k^2x=0$, which was solved on page 23. It is of the simple harmonic or oscillatory type, and we therefore see that the motion of the charge in the primary circuit is oscillatory. Consequently the current is likewise oscillatory, and varies between the limits $+i_0$ and $-i_0$.

The solution of the equation—

$$l_2\frac{di_2}{dt}+m\frac{di_1}{dt}=0$$

is,

$$l_2i_2+mi_1=k,$$

and if $i_1=i_0$, when $i_2=0$—

$$k=mi_0,$$

and,

$$i_2=\frac{m}{l_2}(i_0-i_1).$$

Since i_1 varies between the limits $+i_0$ and $-i_0$, the greatest value of i_2 occurs when $i_1=-i_0$, in which case—

$$i_2=\frac{2mi_0}{l_2}.$$

It will be seen that this is twice the greatest value of the secondary current when the condenser is absent, the drop in the primary current being produced merely by the break at the contact maker. Since the charge passes into the condenser instead of across the gap the sparking at the break is much reduced.

Practical Methods of Measuring Inductances.—The mutual inductance of two coils may be measured by making use of the fact that the quantity of electricity caused to circulate in one when current I is established in the other is $\frac{\text{MI}}{\text{R}}$ coulombs (p. 323). Thus, with the plug in A (Fig. 299) the ballistic throw in the galvanometer G when the current I is started or stopped in the primary circuit may be observed. Then $\frac{\text{MI}}{\text{R}}=\text{Q}=\frac{c\text{T}}{2\pi\text{AH}}\theta$ (see p. 257).

In order to determine $\frac{c}{\text{AH}}$, the plug is removed from A, and two are placed in B and C, which are connected to a very small resistance R_1, say $\frac{1}{100}$ ohm, in the primary circuit. The differ-

ence of potential between the ends of this is now IR_1 and a current $\frac{IR_1}{R}$ flows in the secondary circuit, giving a steady galvanometer deflection θ_1.

Then,
$$\frac{IR_1}{R} \cdot \frac{AH}{c} = \theta_1$$

$$\therefore \frac{MI}{R} = \frac{IR_1}{R\theta_1} \cdot \frac{T}{2\pi}\theta,$$

$$M = \frac{R_1 T}{2\pi\theta_1} \cdot \theta.$$

The time of vibration T of the galvanometer needle may be found in the usual way.

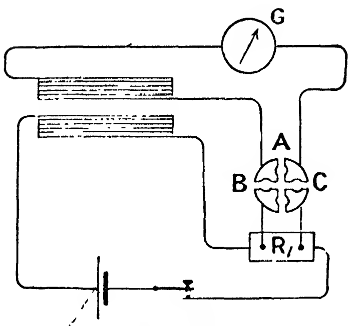


Fig. 299.

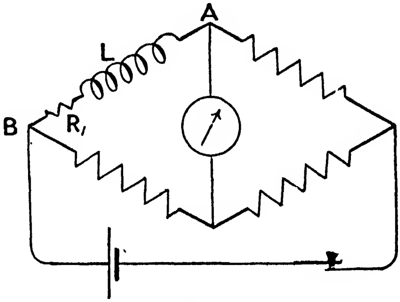


Fig. 300.

Self-Inductance (Rayleigh's Method [1]).—As in the last method, an inductance is measured in terms of a resistance and a time, but here the Wheatstone's bridge is employed. The inductance to be measured is placed in the arm AB of the bridge (Fig. 300) and a balance for steady current obtained in the ordinary way, the battery key being closed before the galvanometer key (not shown in the diagram). On closing the battery key with the galvanometer key already closed, a throw will be obtained, since the balance is disturbed while the current is growing, owing to the extra electromotive force $L\frac{dI}{dt}$ in the arm AB. Any electromotive force in one arm of the bridge causes a proportionate current in every part of the bridge. Let kE be the current in the galvanometer due to electromotive force E in the arm AB.

Then the instantaneous current in the galvanometer due to electromotive force $L\frac{dI}{dt}$ in AB is $kL\frac{dI}{dt}$, and therefore the total

[1] Lord Rayleigh, *Phil. Trans.*, **173**, p. 677. 1882.

quantity of electricity that flows through the galvanometer due to this cause, while the current I_0 is being established in AB is—

$$\int_0^t k \, . \, L\frac{dI}{dt}dt = kL\int_0^{I_0} dI$$
$$= kLI_0.$$
$$\therefore \; kLI_0 = \frac{cT}{2\pi AH}\theta\left(1+\frac{\lambda}{2}\right)$$

where θ is the throw, and λ the logarithmic decrement.

In order to determine $\frac{kI_0AH}{c}$, the resistance in AB is changed by amount R_1, which is so small that there is no appreciable change in the current I_0. The effect is to introduce the small electromotive force I_0R_1 into the arm AB and to produce a steady current kI_0R_1 in the galvanometer. The steady deflection θ_1 produced is given by—

$$kI_0R_1AH = c\theta_1$$
$$\therefore \; \frac{kI_0AH}{c} = \frac{\theta_1}{R_1}$$

and substituting in the above equation, we get—

$$L = \frac{R_1T}{2\pi\theta_1}\theta\left(1+\frac{\lambda}{2}\right).$$

The balance for steady current must be perfect. When the metre bridge is being used this condition is easily attained, but owing to the low resistance of the bridge wire, a Post-Office box, or a suitable combination of resistance boxes, is used by preference. In this case the smallest resistance in the box is usually sufficient to change the steady deflection from one side to the other, and hence a perfect balance cannot be obtained. To get over this difficulty, one of the connections between the boxes may be made with platinoid or manganin wire, and the final adjustment carried out by slipping the wire in the necessary direction through the terminal. The small resistance R_1 may be a standard 0·1, 0·01 or 0·001 ohm, included in the arm AB.

Comparison of Self-Inductances.—The value of a self-inductance in terms of a standard, may be found by placing them one in each adjacent arm of a Wheatstone's bridge, and adjusting the resistances until a balance is obtained for an intermittent, as well as for a steady current. If P, Q, R and S are the respective resistances of the arms of the bridge (Fig. 301), we have, when a balance for steady current is obtained, $\frac{P}{R} = \frac{Q}{S}$.

On closing the galvanometer key and afterwards the battery

key, a galvanometer throw will be observed, unless we have the additional relation $\frac{L_1}{L_2}=\frac{P}{R}=\frac{Q}{S}$.

For the points A and C are at the same potential before the current starts and also when it has become steady, that is, the potential difference between D and A is equal to that between D and C. If now the current grows at the same rate in both branches DAB and DCB, the differences of potential between D and A, and D and C respectively, are equal at every instant, and therefore A and C are always at the same potential, and there will be no current in the galvanometer. The currents grow at the same rate if the time constants of the two circuits are equal, that is—

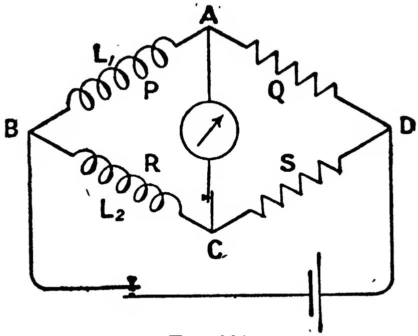


FIG. 301.

$$\frac{L_1}{P+Q}=\frac{L_2}{R+S}, \quad \text{or,} \quad \frac{L_1}{L_2}=\frac{P+Q}{R+S}.$$

But,

$$\frac{P}{R}=\frac{Q}{S}=\frac{P+Q}{R+S};$$

$$\therefore \frac{L_1}{L_2}=\frac{P}{R}=\frac{Q}{S}.$$

Either AB or BC must include a variable resistance, in order that the ratio $\frac{P}{R}$ may be varied. Hence it is necessary to produce a steady balance first; then if the ballistic balance is found to be imperfect, the ratio $\frac{P}{Q}$ must be altered and the process repeated. This is continued until the balance is perfect for both steady and variable currents.

The above methods of measuring inductance are only applicable to a coil which has a non-magnetic core. When there is an iron core a "make" and "break" method must not be used.

In order to increase the sensitiveness to making or breaking the circuit, Ayrton and Perry designed a commutator which they called a Secohmmeter, which on rotation makes the battery, then the galvanometer circuits, then breaks the battery and afterwards

the galvanometer circuits, so that the "break" impulses send a charge through the galvanometer when the ballistic balance is imperfect, but the "make" impulses do not. On reversing the direction of rotation, the charge passes through the galvanometer at the "make" instead of at the "break." The rotation is made rapid by a series of gearing wheels, so that a considerable number of "makes" or "breaks" may be made per second and the galvanometer deflection therefore increased.

Comparison of Capacities (de Sauty).—A method similar to the above may be applied to the comparison of capacities, but in this case the steady current is of course zero. On depressing the key (Fig. 302), a difference of potential is established between B and D and currents flow in the circuits BAD and BED. A and E are at the same potentials at the beginning and after the completion of the charging of the condensers, since the current in both cases is zero. If then the charges on the two condensers have grown at the same rate, A and E are all the time at the same potential and there is no throw of the galvanometer. The charges grow at the same rate when the time constants of the circuits BAD and BED are equal, that is when $R_1C_1 = R_2C_2$ (p. 315).

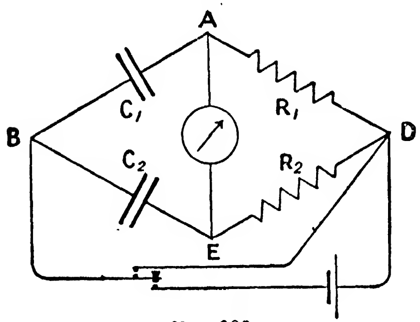


FIG. 302.

Therefore, when the resistances are adjusted until there is no disturbance of the galvanometer on charging or discharging,

$$\frac{C_1}{C_2} = \frac{R_2}{R_1}.$$

When the balance is not attained, the throw is one way on charging by depressing the key, and the other way on discharging by releasing the key.

Circuit with Inductance, Capacity and Resistance (Charge).—We will now find how the current grows in a circuit to which a constant electromotive force e is applied, when the circuit has inductance and capacity as well as resistance. From the cases treated earlier (pp. 308 and 314) we see that the equation of instantaneous electromotive forces will contain four terms and will be,

$$l\frac{di}{dt} + ri + \frac{q}{c} = e,$$

where the letters have the meanings previously assigned to them. Further, $i=\frac{dq}{dt}$, and the equation becomes—

$$l\frac{d^2q}{dt^2}+r\frac{dq}{dt}+\frac{q}{c}=e \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

It will therefore be necessary to solve this for q, afterwards obtaining the current by differentiating the value of q with respect to time.

For convenience we will write, $\frac{r}{l}=2b$, and $\frac{1}{lc}=k^2$, so that—

$$\frac{d^2q}{dt^2}+2b\frac{dq}{dt}+k^2q=\frac{e}{l}.$$

Let $x=q-\frac{e}{lk^2}$, then

$$\frac{dx}{dt}=\frac{dq}{dt}, \quad \text{and,} \quad \frac{d^2x}{dt^2}=\frac{d^2q}{dt^2},$$

and substituting these values in the equation, we get—

$$\frac{d^2x}{dt^2}+2b\frac{dx}{dt}+k^2x=0,$$

which may be solved in the form—

$$x=\epsilon^{\alpha t}.$$

From this we have—

$$\frac{dx}{dt}=a\epsilon^{\alpha t}$$

and,

$$\frac{d^2x}{dt^2}=a^2\epsilon^{\alpha t}$$

so that, substituting these values in the equation, we get—

$$a^2\epsilon^{\alpha t}+2ba\epsilon^{\alpha t}+k^2\epsilon^{\alpha t}=0,$$

or,

$$a^2+2ba+k^2=0$$

that is,

$$a=-b\pm\sqrt{b^2-k^2}$$

and there are two solutions—

$$x=\mathrm{A}'\epsilon^{(-b+\sqrt{b^2-k^2})t}, \text{ and, } x=\mathrm{B}'\epsilon^{(-b-\sqrt{b^2-k^2})t}$$

where A′ and B′ are any arbitrary constants.

Now when two solutions such as—

$$x-\mathrm{A}'\epsilon^{(-b+\sqrt{b^2-k^2})t}=0, \text{ and, } x-\mathrm{B}'\epsilon^{(-b-\sqrt{b^2-k^2})t}=0$$

have been found, it obviously follows that—

$$2x-\mathrm{A}'\epsilon^{(-b+\sqrt{b^2-k^2})t}-\mathrm{B}'\epsilon^{(-b-\sqrt{b^2-k^2})t}=0,$$

or writing A for $\frac{A'}{2}$, and B for $\frac{B'}{2}$,

$$x=A\epsilon^{(-b+\sqrt{b^2-k^2})t}+B\epsilon^{(-b-\sqrt{b^2-k^2})t}$$

or, since $x=q-\frac{e}{k^2l}$,

$$q=A\epsilon^{(-b+\sqrt{b^2-k^2})t}+B\epsilon^{(-b-\sqrt{b^2-k^2})t}+\frac{e}{k^2l}.$$

Remembering that $k^2=\frac{1}{lc}$, we see that the last term is equal to ec. Whatever the nature of the variation of charge may be, we know that after a sufficiently long time has elapsed the charge will become steady and equal to ec (p. 147). Writing this final steady value of the charge equal to q_0, our equation becomes—

$$q=A\epsilon^{(-b+\sqrt{b^2-k^2})t}+B\epsilon^{(-b-\sqrt{b^2-k^2})t}+q_0 \quad . \quad . \quad . \quad \text{(ii)}$$

This equation expresses the mode in which the charge q varies, but it does not give us definitely its value at any given time unless the arbitrary constants A and B are known. These may be determined if we know two conditions as regards charge, or rate of variation of charge, that is, current. Now, in the case considered, the electromotive force is suddenly applied to the circuit, and therefore the charge on the condenser at the instant of application is zero, the expression of which condition is that $q=0$, when $t=0$.

Equation (ii) then reduces to—

$$0=A+B+q_0, \quad \text{or,} \quad A+B=-q_0.$$

Further, the current is zero at the instant of applying the electromotive force, that is, $i=\frac{dq}{dt}=0$, when $t=0$.

Now, from equation (ii)—

$$i=\frac{dq}{dt}=(-b+\sqrt{b^2-k^2})A\epsilon^{(-b+\sqrt{b^2-k^2})t}$$
$$+(-b-\sqrt{b^2-k^2})B\epsilon^{(-b-\sqrt{b^2-k^2})t}$$

$$\therefore\ 0=(-b+\sqrt{b^2-k^2})A+(-b-\sqrt{b^2-k^2})B,$$

or, $$-b(A+B)+\sqrt{b^2-k^2}(A-B)=0,$$

But $A+B=-q_0$,

$$\therefore\ A-B=-\frac{q_0b}{\sqrt{b^2-k^2}}$$

and, $$2A=-q_0\left(1+\frac{b}{\sqrt{b^2-k^2}}\right),\quad 2B=-q_0\left(1-\frac{b}{\sqrt{b^2-k^2}}\right).$$

Putting these values of A and B in equation (ii) we get—

$$q=q_0\left\{1-\frac{1}{2}\left(1+\frac{b}{\sqrt{b^2-k^2}}\right)\epsilon^{(-b+\sqrt{b^2-k^2})t}\right.$$
$$\left.-\frac{1}{2}\left(1-\frac{b}{\sqrt{b^2-k^2}}\right)\epsilon^{(-b-\sqrt{b^2-k^2})t}\right\} \quad . \quad . \quad . \quad \text{(iii)}$$

When $b^2>k^2$, that is $\frac{r^2}{4l^2}>\frac{1}{lc}$, this equation for q cannot be further simplified, and the charge gradually acquires its final value q_0. The mode of approach to q_0 is shown by the dotted curve in Fig. 303, which is drawn for the case in which the

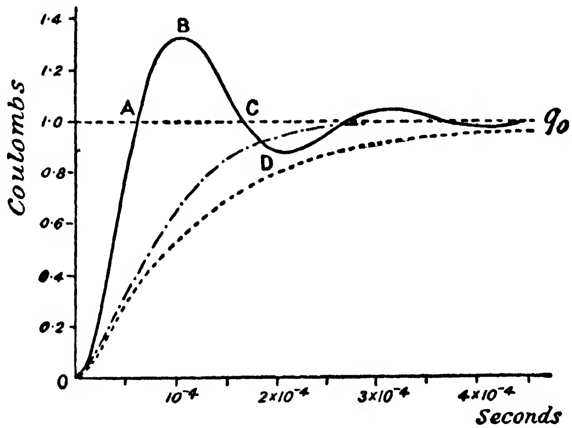


FIG. 303.

inductance is 10 millihenries, the capacity $\frac{1}{10}$ microfarad, and the resistance 1260 ohms.

Thus, $$\frac{R^2}{4L^2}=3{\cdot}97\times10^9$$

and, $$\frac{1}{LC}=10^9,$$

so that, $b=6{\cdot}30\times10^4$, and $k=3{\cdot}16\times10^4$, and if the charging electromotive force is 10^7 volts, the final steady charge Q_0 is 1 coulomb.

If, however, $b<k$ there is an entirely different state of affairs, for b^2-k^2 is negative, and $\sqrt{b^2-k^2}$ is an imaginary quantity. Let it be written, $\sqrt{-1}\sqrt{k^2-b^2}$, or $j\sqrt{k^2-b^2}$, where $j=\sqrt{-1}$,

so that the quantity under the root sign is again real. Equation (iii) thus becomes—

$$q=q_0\left\{1-\epsilon^{-bt}\left(\frac{\epsilon^{j\sqrt{k^2-b^2}t}+\epsilon^{-j\sqrt{k^2-b^2}t}}{2}\right.\right.$$
$$\left.\left.+\frac{b}{\sqrt{k^2-b^2}}\frac{\epsilon^{j\sqrt{k^2-b^2}t}-\epsilon^{-j\sqrt{k^2-b^2}t}}{2j}\right)\right\}$$

Now the exponential form of $\cos\sqrt{k^2-b^2}t$ is $\dfrac{\epsilon^{j\sqrt{k^2-b^2}t}+\epsilon^{-j\sqrt{k^2-b^2}t}}{2}$

and of $\sin\sqrt{k^2-b^2}t$ is $\dfrac{\epsilon^{j\sqrt{k^2-b^2}t}-\epsilon^{-j\sqrt{k^2-b^2}t}}{2j}$ (see p. 377), and substituting these values in our equation we have—

$$q=q_0\left\{1-\frac{\epsilon^{-bt}}{\sqrt{k^2-b^2}}\left(\sqrt{k^2-b^2}\cos\sqrt{k^2-b^2}t+b\sin\sqrt{k^2-b^2}t\right)\right\}.$$

Taking an angle θ, such that—

$$\tan\theta=\frac{b}{\sqrt{k^2-b^2}},\ \sin\theta=\frac{b}{k},\ \text{and}\ \cos\theta=\frac{\sqrt{k^2-b^2}}{k},$$

$$q=q_0\left\{1-\frac{k\epsilon^{-bt}}{\sqrt{k^2-b^2}}\left(\cos\theta\cos\sqrt{k^2-b^2}t+\sin\theta\sin\sqrt{k^2-b^2}t\right)\right\}$$

$$=q_0\left\{1-\frac{k\epsilon^{-bt}}{\sqrt{k^2-b^2}}\cos\left(\sqrt{k^2-b^2}t-\theta\right)\right\}\quad\ldots\ldots\quad\text{(iv)}$$

The full curve OABCD, etc. (Fig. 303), is obtained from this equation, for the case in which $L=10$ millihenries $=10^{-2}$ henries, $C=\frac{1}{10}$ microfarad $=10^{-7}$ farad, and $R=200$ ohms, the charging electromotive force being 10^7 volts, so that $Q_0=EC=1$ coulomb.

$$b^2=\left(\frac{200}{2\times10^{-2}}\right)^2=10^8,\ \text{and},\ k^2=\frac{1}{10^{-2}\times10^{-7}}=10^9.$$

In this case the discharge is oscillatory. The charge is alternately greater and less than Q_0 before settling down to this steady value, and with the given quantities it will be seen that after passing the value Q_0 four times, the amplitude of the oscillation is reduced to about $\frac{1}{40}$ of its original value.

The diminution in amplitude is due to the term ϵ^{-bt}, or $\epsilon^{-\frac{R}{2L}t}$, and with a less resistance than that chosen, this term would become of less importance and the damping of the oscillations less rapid; in the limiting case when $R=0$ the term $\epsilon^{-\frac{R}{2L}t}$ would equal unity, and the charge upon the condenser would vary in a simple harmonic manner.

It will be seen that the value of the charge for the point B upon the curve, that is, after one half-oscillation, is much greater than

its final steady value. This explains the fact that on connecting a condenser to an electrical supply, when the inductance in the circuit is large and the resistance small, the potential difference between the plates momentarily attains nearly twice the final steady value, and the insulation of the condenser may break down, although sufficiently strong to stand the final steady potential difference. The breaking down may be prevented by putting in circuit at the moment of connection, a resistance which may subsequently be removed, but which has the effect of damping out the violent oscillations that would otherwise occur.

Current.—The *current* in the circuit at any instant may be derived from equation (iv); thus—

$$i=\frac{dq}{dt}=q_0k\epsilon^{-bt}\sin(\sqrt{k^2-b^2}t-\theta)+q_0\frac{kb\epsilon^{-bt}}{\sqrt{k^2-b^2}}\cos(\sqrt{k^2-b^2}t-\theta)$$

$$=\frac{q_0k\epsilon^{-bt}}{\sqrt{k^2-b^2}}\{\sqrt{k^2-b^2}\sin(\sqrt{k^2-b^2}t-\theta)+b\cos(\sqrt{k^2-b^2}t-\theta)\}.$$

Remembering that $\tan\theta=\frac{b}{\sqrt{k^2-b^2}}$, etc., we see that this equation becomes—

$$i=\frac{q_0k^2\epsilon^{-bt}}{\sqrt{k^2-b^2}}\sin\sqrt{k^2-b^2}t \quad . \quad . \quad . \quad . \quad \text{(v)}$$

Frequency of Oscillation.—In the case of the current, we see from equation (v) that at times 0, $\frac{\pi}{\sqrt{k^2-b^2}}\cdot\frac{2\pi}{\sqrt{k^2-b^2}}$, etc., the value is zero, and from equation (iv) that after intervals from the start, of $\frac{\theta+\frac{\pi}{2}}{\sqrt{k^2-b^2}}$, $\frac{\theta+\frac{3\pi}{2}}{\sqrt{k^2-b^2}}$, etc., the value of q is equal to q_0.

In either case $\frac{\pi}{\sqrt{k^2-b^2}}$ represents the time for half an oscillation, so that the time for a complete oscillation, or the periodic time, is—

$$\frac{2\pi}{\sqrt{k^2-b^2}}=\frac{2\pi}{\sqrt{\frac{1}{LC}-\frac{R^2}{4L^2}}}.$$

When R is small, this approximates to $2\pi\sqrt{LC}$, and in most practical cases the value of R is not large enough to cause the periodic time to differ greatly from this value.

The frequency is given by—

$$N=\frac{1}{T}=\frac{\sqrt{\frac{1}{LC}-\frac{R^2}{4L^2}}}{2\pi}$$

or when R is small—

$$N=\frac{1}{2\pi\sqrt{LC}}.$$

In the example given on p. 334, $\frac{1}{LC}=10^9$ and $\frac{R^2}{4L^2}=10^8$, so that—

$$N=\frac{\sqrt{10^9-10^8}}{2\pi}=4776 \text{ oscillations per second.}$$

Limiting Case.—Equation (iii) evidently breaks down when $b=k$, for in this case two of the coefficients become infinite. Returning to equation (ii) let us see what form this takes when $\sqrt{b^2-k^2}$ has not vanished but is reduced to some very small quantity h.

Then,
$$\begin{aligned} q&=A\epsilon^{(-b+h)t}+B\epsilon^{(-b-h)t}+q_0\\ &=\epsilon^{-bt}(A\epsilon^{ht}+B\epsilon^{-ht})+q_0 \end{aligned}$$

Writing ϵ^{ht} and ϵ^{-ht} in the form of series—

$$\epsilon^{ht}=1+ht+\frac{h^2t^2}{\underline{|2}}+\frac{h^3t^3}{\underline{|3}}+\ldots$$

$$\epsilon^{-ht}=1-ht+\frac{h^2t^2}{\underline{|2}}-\frac{h^3t^3}{\underline{|3}}+\ldots$$

$$q=\epsilon^{-bt}\left\{A\left(1+ht+\frac{h^2t^2}{\underline{|2}}+\ldots\right)+B\left(1-ht+\frac{h^2t^2}{\underline{|2}}-\ldots\right)\right\}+q_0.$$

Now, h being very small, the terms in h^2 and higher powers may be neglected; but the quantities Ah and Bh have unknown magnitudes since A and B are unknown.

Thus,
$$q=\epsilon^{-bt}\{(A+B)+(A-B)ht\}+q_0.$$

Calling these two constants $(A+B)$ and $(A-B)h$, G and H respectively—

$$q=\epsilon^{-bt}(G+Ht)+q_0.$$

This is a solution of equation (i), as may be shown by differentiation and substitution of q, $\frac{dq}{dt}$ and $\frac{d^2q}{dt^2}$ in (i); it represents the limiting case when $\sqrt{b^2-k^2}$ approaches the value 0.

Now let $q=0$, when $t=0$, and we have $G=-q_0$.

Again, let $i=\frac{dq}{dt}=0$, when $t=0$

$$\frac{dq}{dt}=-b\epsilon^{-bt}(G+Ht)+H\epsilon^{-bt}$$

$$0=-bG+H, \text{ or, } H=-bq_0$$

$$\therefore\ q=q_0\{1-\epsilon^{-bt}(1+bt)\}$$

The curve representing this equation is shown by the chain line in Fig. 303. In this case $L=10^{-2}$ henrys, $C=10^{-7}$ farads, and $R=2\times10^{\frac{5}{2}}$ ohms$=632\cdot4$ ohms, so that—

$$k^2=\frac{1}{LC}=10^9,\ b^2=\frac{R^2}{4L^2}=\frac{4\times10^5}{4\times10^{-4}}=10^9.$$

Discharge.—To find the manner in which the condenser discharges on suddenly removing the external source of electromotive force, we put $e=0$ in equation (i) on p. 331, which then becomes—

$$l\frac{d^2q}{dt^2}+r\frac{dq}{dt}+\frac{q}{c}=0 \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(vi)}$$

or,

$$\frac{d^2q}{dt^2}+2b\frac{dq}{dt}+k^2q=0,$$

where, as before, $b=\frac{r}{2l}$ and $k^2=\frac{1}{lc}$.

This is the same form as the equation in x on p. 331, and, as there explained, the solution is—

$$q=A\epsilon^{(-b+\sqrt{b^2-k^2})t}+B\epsilon^{(-b-\sqrt{b^2-k^2})t} \quad . \quad . \quad . \quad . \quad \text{(vii)}$$

Putting in the limiting conditions we have, since $q=q_0$ when $t=0$,

$$A+B=q_0,$$

and since $\frac{dq}{dt}=0$, when $t=0$—

$$(-b+\sqrt{b^2-k^2})A+(-b-\sqrt{b^2-k^2})B=0.$$

Solving these two equations for A and B and substituting their values in (vii)—

$$q=q_0\left\{\tfrac{1}{2}\left(1+\frac{b}{\sqrt{b^2-k^2}}\right)\epsilon^{(-b+\sqrt{b^2-k^2})t}+\tfrac{1}{2}\left(1-\frac{b}{\sqrt{b^2-k^2}}\right)\epsilon^{(-b-\sqrt{b^2-k^2})t}\right\} \quad \text{(viii)}$$

When $b>k$ this equation represents the dead beat discharge, the dotted curve in Fig. 304 being drawn for the case given on p. 333. When $b<k$ the equation is transformed as on p. 334, into the form—

$$q=q_0\epsilon^{-bt}\left\{\frac{\epsilon^{j\sqrt{k^2-b^2}t}+\epsilon^{-j\sqrt{k^2-b^2}t}}{2}+\frac{b}{\sqrt{k^2-b^2}}\,\frac{\epsilon^{j\sqrt{k^2-b^2}t}-\epsilon^{-j\sqrt{k^2-b^2}t}}{2j}\right\}$$

or,

$$q=q_0\epsilon^{-bt}\left\{\cos\sqrt{k^2-b^2}t+\frac{b}{\sqrt{k^2-b^2}}\sin\sqrt{k^2-b^2}t\right\}$$

$$=\frac{q_0k\epsilon^{-bt}}{\sqrt{k^2-b^2}}\cos(\sqrt{k^2-b^2}t-\theta) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ix)}$$

where $\tan\theta=\frac{b}{\sqrt{k^2-b^2}}$, etc.

The discharge is therefore oscillatory when $k>b$, and the curve q_0ABD in Fig. 304 shows the form of the discharge in the given case. At B the charge upon the condenser has the reverse sign to that of q_0. The points where the curve cuts the axis are separated by an interval of time $\frac{\pi}{\sqrt{k^2-b^2}}$ or $\frac{\pi}{\sqrt{\frac{1}{LC}-\frac{R^2}{4L^2}}}$, which is therefore half the periodic time.

The oscillatory character of the discharge of a condenser when

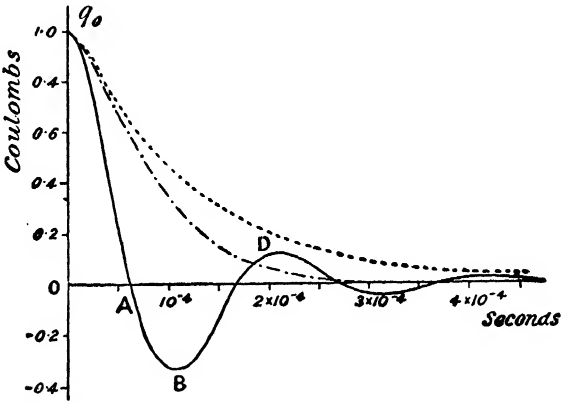


FIG. 304.

the inductance in the circuit is sufficiently great for the condition $\frac{R^2}{4L^2}<\frac{1}{LC}$ to be fulfilled, was first pointed out by Lord Kelvin in 1853.[1]

In the event of R being inappreciable, $b=0$, and the equation for the discharge becomes—

$$Q=Q_0 \cos \frac{t}{\sqrt{LC}},$$

and the oscillations are harmonic, the charge surging backwards and forwards through the inductance, being alternately positive and negative upon either plate of the condenser. For either extreme, the energy of the charge is $\frac{1}{2}\frac{Q_0^2}{C}$, and at an instant half-way between these extremes, the charge upon the condenser is zero, and the energy is associated with the inductance and has the value $\frac{1}{2}LI^2$ joules (see p. 306).

Since—
$$I=\frac{dQ}{dt}=-\frac{1}{\sqrt{LC}}Q_0 \sin \frac{t}{\sqrt{LC}},$$

[1] W. Thomson, *Phil. Mag.* (Ser. 4), 5, p. 393. 1853.

the greatest value of this occurs when—

$$\frac{t}{\sqrt{LC}}=\frac{\pi}{2}, \text{ and, } \sin\frac{t}{\sqrt{LC}}=1.$$

Hence at this instant, $I=-\dfrac{Q_0}{\sqrt{LC}}$

and the energy, $\frac{1}{2}LI^2=\frac{1}{2}L\,.\,\dfrac{Q_0^2}{LC}=\frac{1}{2}\dfrac{Q_0^2}{C}.$

Thus the energy alternates between the statical and the dynamical form, and is associated first with the capacity and then with the inductance, being reversed twice at every oscillation.

When R is not zero, there is a continual dissipation of electrical energy into heat as the current flows in the resistance, and since this process is not reversible, the energy of the charge gets less and less at each oscillation and finally vanishes. It is also possible that some of the energy is radiated into space at each oscillation, but a discussion of this will be reserved to a later chapter.

Limiting Conditions.—The distinction between the dead beat form of discharge, in which the sign of the charge is not reversed ($b>k$), and the oscillatory form ($b<k$), although very much in evidence in the mathematical form in which the discharge is represented, is not really very marked, for as k gets nearer and nearer in value to b, the oscillations become so rapidly damped that the discharge is almost exponential in type. In fact there is no discontinuity in passing from one form of the discharge to the other, as the student may see if he will plot several curves resembling those in Fig. 304, for values of k approaching very near to that of b.

As before, equation (viii) breaks down when $b=k$, and we obtain our solution by putting $\sqrt{b^2-k^2}=h$, and finding the form of the solution when h is extremely small On p. 336 we saw that the exponential terms take the form $\epsilon^{-bt}(G+Ht)$, and therefore—

$$q=\epsilon^{-bt}(G+Ht).$$

If now, $q=q_0$, when $t=0$, $G=q_0$; and if $\dfrac{dq}{dt}=0$, when $t=0$, $H=bq_0$,

$$\therefore\ q=q_0\epsilon^{-bt}(1+bt).$$

The chain curve in Fig. 304 represents this equation in the given case.

Rate of Discharge.—When b is very small in comparison with k, the oscillations are nearly simple harmonic, and die away very slowly; and on the other hand, when b is very large in com-

parison with k, the discharge is dead beat, the steady state being reached only after the lapse of considerable time. Thus when b is very small or very large with respect to k, the process of discharge takes considerable time for its completion. This consideration, together with an examination of Fig. 304, will lead us to the conclusion that the final state of complete discharge is reached more rapidly as b approaches in value to k.

From equation (ix) (p. 337) we see that the term

$$\cos(\sqrt{k^2-b^2}t-\theta)$$

has alternately values $+1$ and -1, the values $+1$ occurring at times differing by intervals $\frac{2\pi}{\sqrt{k^2-b^2}}$; hence the maxima occur at times differing by these intervals, and have values $q \propto q_0 \epsilon^{-\frac{2\pi nb}{\sqrt{k^2-b^2}}}$, where n has the successive values, 0, 1, 2, 3, etc. Remembering that for the oscillatory discharge, $b<k$, we see that these successive maxima decrease most rapidly as b increases, since $\frac{b}{\sqrt{k^2-b^2}}$ becomes greater. We have the maximum rate of damping out of the vibrations when $b=k$, since in this limiting condition—

$$\frac{b}{\sqrt{k^2-b^2}}=\infty.$$

Returning to equation (vii) (p. 337) we see that for the dead beat discharge, that is, when $b>k$; the second term diminishes extremely rapidly since $-b-\sqrt{k^2-b^2}$ is always numerically large. The first term is therefore much the more important, and the greater the numerical value of the term $-b+\sqrt{b^2-k^2}$, which is always negative, the more rapidly does the charge decay. Thus, as $b-\sqrt{b^2-k^2}$ increases the decay becomes more rapid. To find how $b-\sqrt{b^2-k^2}$ varies as b approaches k in value, write—

$$y=b-\sqrt{b^2-k^2}$$

$$\frac{dy}{db}=1-\frac{b}{\sqrt{b^2-k^2}},$$

and as b approaches in value to k, $\frac{dy}{db}$ is evidently negative, that is, as b diminishes $b-\sqrt{b^2-k^2}$ increases, and therefore as b diminishes the rate of decay increases. This holds up to the limiting case when $b=k$, and we therefore see that as we approach from the oscillatory side or the dead beat side, towards the limiting case, the rate of discharge becomes more rapid. Hence

the limiting case when $b=k$ is that in which the charge most quickly disappears.

This was first pointed out by Dr. Sumpner [1] and is of particular importance in the design of galvanometers. There is complete analogy between the motion of a galvanometer needle and that of a discharging condenser (see p. 262). With no damping, the swings last a great time, as in the case of the ballistic galvanometer. On the other hand, if the damping is too great the motion is dead beat, and the final position of the needle is taken up too slowly. This is extremely troublesome as it renders the galvanometer very sluggish. The best working conditions for ordinary use are obtained when the needle comes very quickly to rest after very few oscillations, a condition sometimes, although wrongly, called dead beat.

The term *critical damping* is often applied to the case corresponding to $b=k$ in the above and it is nowadays usual for the maker to state the critical damping resistance of an aperiodic moving-coil galvanometer. This enables the user to match a circuit to the galvanometer if time spent in observing the deflections is to be economised.

Discharge examined by Rotating Mirror.—The change in character of the discharge from dead beat to oscillatory, as the resistance of the circuit is reduced, may be seen by examining the spark by means of a rotating mirror. The condenser is charged by means of the induction coil, the spark discharge taking place between the terminals P (Fig. 305) once at every break of the primary circuit. The circuit ACBP in which the oscillations occur, has usually sufficient inductance for the purpose, but if necessary this may be increased by including a few turns of wire. When the points P are far apart, the image of the spark seen in the rotating mirror consists of one thin band for each discharge, but on approaching the points until the resistance falls below the critical value, a group of lines which may extend to 6 or 7 in number, is seen at each discharge, which is consequently oscillatory.

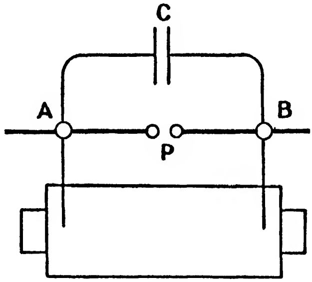


FIG. 305.

The production of oscillations will be further considered in the chapter on radiation.

[1] W. E. Sumpner, *Phil. Mag.* (Ser. 5), 25, p. 453. 1888.

CHAPTER XI

ALTERNATING CURRENTS

THE development of dynamo-electric machinery, in which a coil or system of coils is rotated in, or moves through, a magnetic field, giving rise to alternating electromotive forces, and of the transformer which enables the energy produced to be changed from small current at high voltage to large current at low voltage and *vice versâ*, has rendered a study of the current produced by an electromotive force which varies harmonically, of great importance. A still further importance to this study arises from the employment of oscillating currents, the commercial application of which is realised in the various systems of wireless telegraphy and telephony.

Circuit with Inductance and Resistance.—For many purposes, the capacity in the circuit is relatively unimportant, and we will first consider the simple case in which a harmonic electromotive force, $E_0 \sin pt$, is applied to a circuit having resistance and inductance only, the effect of including capacity as well being left until later.

The equation of electromotive forces for such a circuit has been given on p. 307. It is—

$$L\frac{dI}{dt}+RI=E.$$

Replacing E by the value $E_0 \sin pt$, we have—

$$L\frac{dI}{dt}+RI=E_0 \sin pt,$$

E_0 being the maximum electromotive force, and $\frac{2\pi}{p}$ the periodic time of alternation.

A pure sine electromotive force, that is, one which varies in a simple harmonic manner, may be produced by rotating a coil with constant angular velocity in a uniform magnetic field. For, if N be the magnetic flux linked with the coil when its plane is perpendicular to the field (Fig. 249), N sin θ is the flux when the plane of the coil makes angle θ with the field.

Then,
$$E=-\frac{d(N\sin\theta)}{dt}$$
$$=-N\frac{d(\sin\theta)}{dt}$$
$$=-N\cos\theta\,.\,\frac{d\theta}{dt}.$$

But if the angular velocity $\frac{d\theta}{dt}=p$, then, $d\theta=pdt$, or $\theta=pt+\frac{\pi}{2}$, the constant of integration being $\frac{\pi}{2}$, if time is reckoned from the instant at which $\theta=+\frac{\pi}{2}$.

Then,
$$E=-pN\cos\left(pt+\frac{\pi}{2}\right)$$
$$=pN\sin pt$$
$$=E_0\sin pt.$$

The electromotive force produced by an alternating current dynamo is not usually of this simple type, but is of the form—

$$E_0\sin pt+E_1\sin 3pt+E_2\sin 5pt+\ldots,$$

but we shall only deal with the case in which the first term alone is present, this term being always by far the most important.

The only part of the solution of our equation which is of importance to us is that in which the current has the same periodicity as the electromotive force, any other being quickly damped out. The current after a very short time takes the form, $I=A\sin pt+B\cos pt$, which is the most general form of a simple harmonic current of periodic time $\frac{2\pi}{p}$. A and B are constants to be determined from the conditions of the problem. To find them, differentiate I with respect to t, and substitute for I and $\frac{dI}{dt}$ in the original equation.

Thus,
$$\frac{dI}{dt}=pA\cos pt-pB\sin pt,$$

and substituting in the equation on p. 342, we have—

$$LpA\cos pt-LpB\sin pt+RA\sin pt+RB\cos pt=E_0\sin pt.$$

This must be true for all values of t.

When, $t=0$, $\sin pt=0$, and $\cos pt=1$.

$$\therefore LpA+RB=0.$$

When, $pt=\frac{\pi}{2}$, $\sin pt=1$, and $\cos pt=0$.

$$\therefore \text{RA}-\text{L}p\text{B}=\text{E}_0.$$

Solving these two simultaneous equations for A and B, we get—

$$\text{A}=\frac{\text{RE}_0}{\text{L}^2p^2+\text{R}^2}, \text{ and, } \text{B}=-\frac{\text{L}pE_0}{\text{L}^2p^2+\text{R}^2}.$$

Hence, $$\text{I}=\frac{\text{E}_0}{\text{L}^2p^2+\text{R}^2}(\text{R}\sin pt-\text{L}p\cos pt).$$

This may be thrown into a more convenient form by writing $\cos\theta$ for $\frac{\text{R}}{\sqrt{\text{L}^2p^2+\text{R}^2}}$, whence—

$$\frac{\text{L}p}{\sqrt{\text{L}^2p^2+\text{R}^2}}=\sin\theta, \text{ and, } \frac{\text{L}p}{\text{R}}=\tan\theta.$$

$$\therefore \text{I}=\frac{\text{E}_0}{\sqrt{\text{L}^2p^2+\text{R}^2}}(\sin pt\cos\theta-\cos pt\sin\theta),$$

$$\text{I}=\frac{\text{E}_0}{\sqrt{\text{L}^2p^2+\text{R}^2}}\sin(pt-\theta).$$

We therefore see that the current has amplitude $\frac{\text{E}_0}{\sqrt{\text{L}^2p^2+\text{R}^2}}=\text{I}_0$ and lags in phase behind the electromotive force by an angle

$$\theta=\tan^{-1}\frac{\text{L}p}{\text{R}}.$$

In the special case in which $\text{L}=0$—

$$\text{I}=\frac{\text{E}_0\sin pt}{\text{R}}=\frac{\text{E}}{\text{R}}.$$

The current is then in phase with the electromotive force, and its value at every instant is given by the ordinary ohmic relation.

Vector Diagram.—The behaviour of a circuit to which an alternating electromotive force is applied, may conveniently be represented by a vector diagram of electromotive forces, and by means of such a diagram the current may be plotted in the form of a curve. Thus if OA in Fig. 306 (i) represents to scale the value of E_0, then the projection of this on the axis of y at the instant t, is $\text{E}_0\sin pt$ and is equal to E, the instantaneous value of the electromotive force. On constructing the right-angled triangle AOB with angle $\text{AOB}=\theta=\tan^{-1}\frac{\text{L}p}{\text{R}}$,

$$\text{OB}=\text{E}_0\cos\theta=\frac{\text{RE}_0}{\sqrt{\text{L}^2p^2+\text{R}^2}}=\text{RI}_0,$$

and OD, the projection of OB upon the axis of y, is—

$$\frac{RE_0}{\sqrt{L^2p^2+R^2}}\sin(pt-\theta)=RI_0\sin(pt-\theta)=RI,$$

and is the instantaneous value of that component of the applied electromotive force that is used in driving the current in opposition to the ohmic resistance of the circuit.

Again, since $DE=OE-OD=E-RI$, and $L\frac{dI}{dt}=E-RI$, we see that DE represents the component of the electromotive force

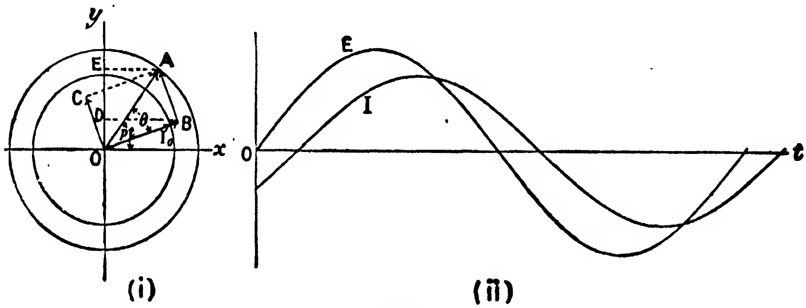


FIG. 306.

that is due to the inductance of the circuit, and is the projection of BA upon the axis of y. But $BA=OA\sin\theta=\frac{LpE_0}{\sqrt{L^2p^2+R^2}}$ $=LpI_0$. BA is at right angles to OB, and the electromotive force due to inductance is therefore 90° in phase ahead of the current.

This might also have been obtained from the relation—

$$I=I_0\sin(pt-\theta),$$

for,
$$\frac{dI}{dt}=pI_0\cos(pt-\theta)=pI_0\sin(pt-\theta+90^\circ).$$

$$\therefore L\frac{dI}{dt}=LpI_0\sin(pt-\theta+90^\circ).$$

The vector OC may be drawn parallel and equal to BA, and we see that the angle $xOC=pt-\theta+90^\circ$, and electromotive force OA is resultant of OB and OC. As the three rotate with constant angular velocity p, their projections on the axis of y represent the instantaneous values of the applied electromotive force and its components. The current being in phase with OB, it may be represented by a vector $OI_0=\frac{OB}{R}$. Then by taking an axis of time parallel to Ox we may, by making the ordinates

equal to OE and to the projection of OI_0, draw the curves of electromotive force (E) and current (I) in Fig. 306 (ii).

Impedance and Reactance.—The quantity $\sqrt{L^2p^2+R^2}$ plays a similar part in the consideration of alternating currents, to resistance in continuous currents, for $I_0=\frac{E_0}{\sqrt{L^2p^2+R^2}}$. It is called the *Impedance* of the circuit. If either L or p becomes zero, the impedance reduces to R, the resistance. The greater the value of Lp, the more does the impedance differ from the resistance, and when the periodicity n $\left(\text{that is } \frac{p}{2\pi}\right)$ becomes so great that the resistance is relatively unimportant, the impedance becomes Lp. This quantity is called the *Reactance* of the circuit, and the impedance has for its limiting values the resistance, for very low frequency, and the reactance, for very high frequency.

Measuring Instruments.—An ordinary electromagnetic ammeter or voltmeter whose moving parts are comparatively massive, will indicate the mean value of the quantity to be measured, when this varies rapidly. The mean value of the quantity $E_0 \sin a$ for a complete cycle, where a is written for simplicity instead of pt, is—

$$\frac{\int_0^{2\pi} E_0 \sin a \, da}{\int_0^{2\pi} da} = -\frac{E_0}{2\pi}\Big[\cos a\Big]_0^{2\pi} = 0,$$

and therefore the reading of an electromagnetic voltmeter on an alternating supply will be zero. The reason is, that for the first half of a cycle, the mean value is $\frac{1}{\pi}\int_0^{\pi} E_0 \sin a \, da = -\frac{E_0}{\pi}\Big[\cos a\Big]_0^{\pi} = \frac{2E_0}{\pi}$, and for the second half, $\frac{1}{\pi}\int_{\pi}^{2\pi} E_0 \sin a \, da = -\frac{2E_0}{\pi}$. Thus the mean values are alternately positive and negative, and the suspended needle or coil will receive equal and opposite impulses during a complete cycle. For this reason it is necessary to employ instruments whose deflections are all in one direction, whatever the direction of the electromotive force or current. This is the case when the deflection depends upon the square of the electromotive force or current, as in the case of the hot-wire voltmeter or ammeter, the heating produced by the current in a fine wire causing its expansion, which the instrument indicates. Also the electrometer used idiostatically (see p. 158) may be used for the measurement of alternating electromotive forces; or certain soft-iron ammeters (p. 83) in the case of current.

In order to interpret the readings of such an instrument, let us find the mean value of $I_0^2 \sin^2 a$, or $E_0^2 \sin^2 a$, for it is this mean value that is proportional to the indication of the instrument; that is, we must find the value of the quantity, $\dfrac{\int_0^{2\pi} I_0^2 \sin^2 a\, da}{\int_0^{2\pi} da}$,

Now,
$$\int_0^{2\pi} da = 2\pi,$$

and,
$$\int_0^{2\pi} I_0^2 \sin^2 a da = I_0^2 \int_0^{2\pi} \frac{1-\cos 2a}{2} da$$
$$= I_0^2 \left[\frac{a}{2} - \frac{\sin 2a}{4}\right]_0^{2\pi}$$
$$= \pi I_0^2.$$

$$\therefore \text{ mean value} = \frac{\pi I_0^2}{2\pi} = \frac{I_0^2}{2}.$$

Virtual Current and E.M.F.—The same value of the mean would have been obtained if we had taken half the cycle instead of the whole, since the square of any quantity is positive, and the sign of $I_0^2 \sin^2 a$ does not change during the cycle. In Fig. 307 the

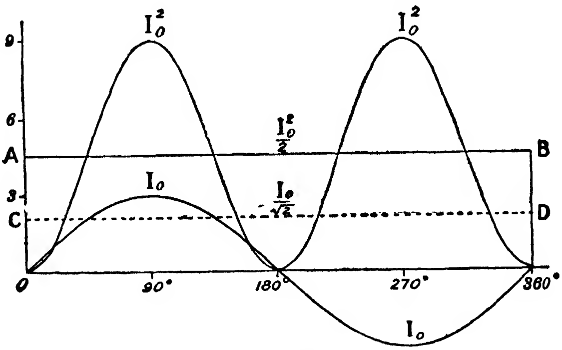


FIG. 307.

values of I and I^2 are plotted, when $I_0=3$. The line AB has the same mean ordinate as the curve I^2. The continuous current represented by CD, whose square would have the same mean value as that of the alternating current is therefore $\frac{I_0}{\sqrt{2}}$, and would give the same reading on a hot-wire ammeter. This is called the *virtual current*, and is equivalent to the alternating

current whose maximum value is I_0. Thus if $I_0 \sin a$ is measured in amperes, $\frac{I_0}{\sqrt{2}}$ is the equivalent current measured in virtual amperes.

In a similar way $\frac{E_0}{\sqrt{2}}$ is the virtual voltage of an alternating current whose maximum voltage is E_0.

The advantage of measuring electromotive force in virtual volts, and current in virtual amperes, is that a suitable instrument may be calibrated by means of a continuous electromotive force or current, and will then read virtual volts or amperes on an alternating current supply.

Measurement of Inductance.—When a source of alternating electromotive force is available, the self-inductance of a coil may be found by measuring the electromotive force and current, first using a direct, and then an alternating current. The first readings give the resistance R of the coil. The second readings give $\frac{I_0}{\sqrt{2}}$ and $\frac{E_0}{\sqrt{2}}$.

But,
$$\frac{I_0}{\sqrt{2}} = \frac{E_0}{\sqrt{2}\sqrt{L^2p^2+R^2}}.$$

$$\therefore\ L^2p^2+R^2 = \frac{E_0^2}{I_0^2} = \frac{E_\sim^2}{I_\sim^2},$$

where $E_\sim$ and $I_\sim$ are the virtual volts and amperes.

If the frequency of alternation is $n_\sim$,

$$n_\sim = \frac{p}{2\pi}, \quad \therefore p = 2\pi n_\sim,$$

and L may then be calculated.

This method has the advantage that when the circuit encloses iron, and the inductance is therefore variable, the value obtained is that for the particular value of $I_\sim$ employed, and this may be chosen to be the value at which the coil is eventually to be used.

Power in Alternating Current Circuit.—In the case of a non-inductive circuit, $I = \frac{E_0}{R} \sin pt = I_0 \sin pt$, when $E = E_0 \sin pt$. The rate of working at any instant is therefore $EI = E_0 I_0 \sin^2 pt$ watts.

We have seen (p. 347) that the mean value of $\sin^2 pt$ for a cycle is equal to $\frac{1}{2}$, and therefore—

$$\text{mean rate of working} = \tfrac{1}{2}E_0 I_0 \text{ watts}$$
$$= \frac{E_0}{\sqrt{2}} \cdot \frac{I_0}{\sqrt{2}} \text{ watts.}$$

Now $\frac{E_0}{\sqrt{2}}$ is the virtual voltage, and $\frac{I_0}{\sqrt{2}}$ the virtual current, therefore mean rate of working in watts is the product (virtual volts) × (virtual amperes).

In the case of an inductive circuit—

$$E = E_0 \sin pt,$$
$$I = I_0 \sin (pt - \theta).$$

Therefore the instantaneous rate of working is—

$$\begin{aligned} EI &= E_0 I_0 \sin pt \sin (pt - \theta) \\ &= E_0 I_0 \sin pt (\sin pt \cos \theta - \cos pt \sin \theta) \\ &= E_0 I_0 \sin^2 pt \cos \theta - \tfrac{1}{2} E_0 I_0 \sin 2pt \sin \theta. \end{aligned}$$

The mean value for a cycle is $\frac{1}{2}$ in the case of $\sin^2 pt$, and zero for $\sin 2pt$.

$$\begin{aligned} \therefore \text{ mean rate of working} &= \tfrac{1}{2} E_0 I_0 \cos \theta \text{ watts} \\ &= (\text{virtual volts})(\text{virtual amperes}) \times \cos \theta. \end{aligned}$$

The following method brings us to the same result, and throws additional light upon the processes going on.

Let the vector E_0 (Fig. 308) represent the maximum electromotive force, and I_0 the maximum current. Resolving I_0 along and at right angles to E_0, we get the component $I_0 \cos \theta$ in phase with the electromotive force, giving us the mean rate of working $\frac{1}{2} E_0 I_0 \cos \theta$, and the component $I_0 \sin \theta$, lagging 90° in phase behind the electromotive force, and giving us a mean rate of working

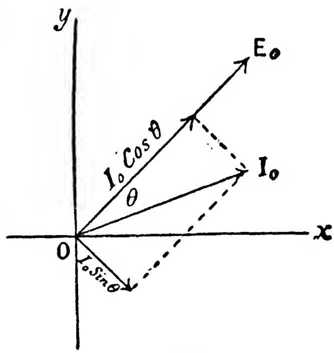


FIG. 308.

$$\frac{1}{2\pi}\int_0^{2\pi} E_0 I_0 \sin a \sin(a - 90°) da,$$

that is—

$$-\frac{E_0 I_0}{2\pi}\int_0^{2\pi} \sin a \cos a \, da = \frac{E_0 I_0}{8\pi}\Big[\cos 2a\Big]_0^{2\pi} = 0.$$

This latter component is called *Idle* or *Wattless* current since its presence does not contribute to the rate at which work is being done in the circuit. When the inductance of a circuit is so great in comparison with the resistance that the latter may be neglected, the current is entirely wattless; for in this case, $\theta = 90°$, and $\cos \theta = 0$.

The curves for E and I in Fig. 309 are drawn for a case in which the current lags behind the electromotive force by an amount *ac*. For the parts of the cycle *ce* and *fg* the electro-

motive force and current are in the same direction, and the source of energy is doing work upon the circuit. If we call this work positive, then for the parts *ac* and *ef* of the cycle, when E and I are in opposite directions, the work is negative, which means that the electromotive force is employed in diminishing the current, or the circuit is doing work upon the source of energy. The curve *abcde*, etc., is drawn by taking ordinates at each point equal to the product of E and I, and the large loops such as *cde* represent work done on the circuit, and the small loops such as *abc* represent energy drawn from it. The difference in area of the two is a measure of the work done in one half cycle.

When $L=0$, $\theta=0$, and the current is in phase with the electromotive force; the small loops are then absent. But as L increases, the difference of phase increases up to 90° in the limiting case when R is negligible, and in this case the positive and

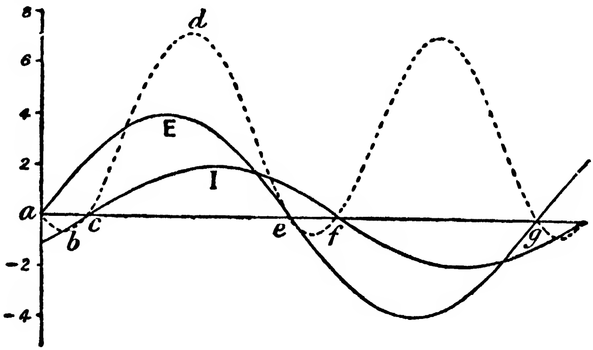


FIG. 309.

negative loops are equal in size. Their difference is then zero, and the current is entirely wattless.

The case is similar to that of a frictionless pendulum ; although the motion is alternating, the total work done by gravity upon the pendulum in a cycle is zero.

Power Factor.—On measuring separately the virtual volts and virtual amperes for a given circuit by means of a voltmeter and an ammeter and taking the product, we obtain the apparent watts. This is not the actual power absorbed in the circuit, unless the current is in phase with the electromotive force, for the product has yet to be multiplied by $\cos\theta$ to obtain the true watts. The ratio of true watts to apparent watts is called the power factor of the circuit.

Since,—true watts=(apparent watts)$\times\cos\theta$, we see that the power factor is equal to $\cos\theta$, where $\theta=\tan^{-1}\frac{Lp}{R}$.

Wattmeters.—The true power absorbed in any circuit may be

measured directly by means of a suitable wattmeter, whose low resistance coil P is in series with the circuit, the high resistance shunt coil Q being connected to the points A and B (Fig. 310) between which the power to be measured is absorbed. In the case of the Kelvin watt balance (p. 248) the shunt circuit has in it a non-inductive resistance R_1 of high value. The mechanical force or couple between the two coils is proportional to the product of the currents in them, that in P being the current I flowing in AB, and that in Q being proportional to, and in phase with, E the difference of potential between A and B. The instrument is calibrated to read directly in watts, and will therefore give the mean value of EI, or the true power absorbed between A and B.

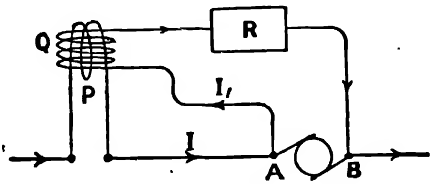


FIG. 310.

A difficulty arises owing to the fact that the coil Q always has some inductance, and the current in it therefore lags behind the electromotive force in it, causing the indicated mean value of the watts to be too great. In the Kelvin instrument this lag is reduced to a negligible amount by making the non-inductive resistance R_1 in the shunt circuit very great; while in the Addenbrooke electrostatic wattmeter, which is a modified quadrant electrometer, there is no appreciable lag.

If I_1 be the current in the shunt circuit Q, L_1 being the inductance, and R the resistance of the shunt circuit—

$$I_1 = \frac{E_0}{\sqrt{L_1^2 p^2 + R^2}} \sin (pt - \theta_1).$$

Also the current in the main circuit is—

$$I = I_0 \sin (pt - \theta),$$

and the instrument indicates the mean value of the product $I \times I_1$, that is—

$$\frac{E_0 I_0 \sin (pt - \theta) \sin (pt - \theta_1)}{\sqrt{L_1^2 p^2 + R^2}}.$$

Expanding the terms $\sin (pt - \theta)$ and $\sin (pt - \theta_1)$, and taking the product, this quantity becomes—

$$\frac{E_0 I_0}{\sqrt{L_1^2 p^2 + R^2}} \{\sin^2 pt \cos \theta \cos \theta_1 + \cos^2 pt \sin \theta \sin \theta_1 \\ - \tfrac{1}{2} \sin 2pt \sin (\theta + \theta_1)\}.$$

The mean values of $\sin^2 pt$ and $\cos^2 pt$ are both $\frac{1}{2}$, and of $\sin 2pt$ is zero, and the mean of the whole expression is therefore—

$$\frac{1}{2}\frac{I_0E_0}{\sqrt{L_1{}^2p^2+R^2}}\cos(\theta-\theta_1).$$

If $L_1=0$, the true power $\frac{1}{2}\frac{I_0E_0}{R}\cos\theta$ would be indicated.

$$\therefore \frac{\text{indicated power}}{\text{true power}}=\frac{R}{\sqrt{L_1{}^2p^2+R^2}}\cdot\frac{\cos(\theta-\theta_1)}{\cos\theta}$$

$$=\frac{\cos\theta_1\cos(\theta-\theta_1)}{\cos\theta}$$

$$=\frac{\cos\theta_1(\cos\theta\cos\theta_1+\sin\theta\sin\theta_1)}{\cos\theta}$$

$$=\cos^2\theta_1+\tfrac{1}{2}\tan\theta\sin 2\,\theta_1.$$

Since L_1 is small, this ratio approaches the value unity for $L_1=0$. On differentiating it with respect to θ_1 it may be seen that the rate of change is essentially positive, so that for values of θ_1 slightly greater than 0 the ratio is greater than unity, and the instrument reads too high.

The power factor of a circuit may be measured by finding the true watts absorbed in it, by means of a wattmeter, and the apparent watts by means of an ammeter and voltmeter.

$$\text{Then, power factor}=\frac{\text{true watts}}{\text{apparent watts}}=\cos\theta.$$

Circuit containing Capacity, Inductance and Resistance.—The equation of electromotive forces for a circuit having capacity, inductance and resistance is (see p. 330)

$$L\frac{dI}{dt}+RI+\frac{Q}{C}=E_0\sin pt,$$

or, since $I=\frac{dQ}{dt}$—

$$L\frac{d^2Q}{dt^2}+R\frac{dQ}{dt}+\frac{Q}{C}=E_0\sin pt.$$

The simple harmonic part of the solution of this equation, that is, $Q=A\sin pt+B\cos pt$, may be found in an exactly similar manner to that given on p. 343. By differentiating, substituting, and solving the simultaneous equations for A and B, we find that—

$$A=\frac{E_0\left(\frac{1}{Cp}-Lp\right)}{p\left\{\left(\frac{1}{Cp}-Lp\right)^2+R^2\right\}},\text{ and, }B=-\frac{E_0R}{p\left\{\left(\frac{1}{Cp}-Lp\right)^2+R^2\right\}},$$

thus—

$$Q=\frac{E_0\left(\frac{1}{Cp}-Lp\right)}{p\left\{\left(\frac{1}{Cp}-Lp\right)^2+R^2\right\}}\sin pt-\frac{E_0R}{p\left(\left\{\frac{1}{Cp}-Lp\right)^2+R^2\right\}}\cos pt,$$

and further—

$$I=\frac{dQ}{dt}=\frac{E_0\left(\frac{1}{Cp}-Lp\right)}{\left\{\left(\frac{1}{Cp}-Lp\right)^2+R^2\right\}}\cos pt+\frac{E_0R}{\left\{\left(\frac{1}{Cp}-Lp\right)^2+R^2\right\}}\sin pt.$$

Taking as before, $\tan\theta=\dfrac{\frac{1}{Cp}-Lp}{R}$,

we have, $$I=\frac{E_0}{\sqrt{\left(\frac{1}{Cp}-Lp\right)^2+R^2}}\sin(pt+\theta).$$

The impedance is in this case $\sqrt{\left(\frac{1}{Cp}-Lp\right)^2+R^2}$.

The following four special cases are of interest.

(i) When $L=0$ and $C=\infty$, then $\theta=0$ and $\tan\theta=0$, and the equation reduces to $I=\frac{E_0}{R}\sin pt$.

This case has already been discussed (p. 344).

(ii) When $C=\infty$, then $\tan\theta=-\frac{Lp}{R}$, and the equation is—

$$I=\frac{E_0}{\sqrt{L^2p^2+R^2}}\sin(pt-\theta).$$

The vector diagram and the electromotive force and current curves are given in Fig. 306.

(iii) When $L=0$, then $\tan\theta=\frac{1}{CpR}$,

and, $$I=\frac{E_0}{\sqrt{\frac{1}{C^2p^2}+R^2}}\sin(pt+\theta).$$

The current is here in advance of the electromotive force by the difference in phase, $\theta=\tan^{-1}\frac{1}{CpR}$.

The vector diagram is given in Fig. 311 (i) and the electromotive force and current curves for one cycle in Fig. 311 (ii).

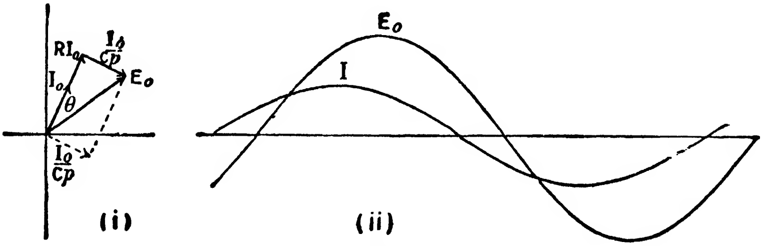


FIG. 311.

The power absorbed in the circuit is again $\frac{1}{2}I_0E_0 \cos \theta$, and the idle or wattless component of the current is $I_0 \sin \theta$ (see p. 349).

If in addition, $R=0$, then $\theta=90°$, and the current is entirely wattless.

(iv) When $$Lp=\frac{1}{Cp}, \ \theta=0,$$

then, $$I=\frac{E_0}{R} \sin pt.$$

In this case the current is in phase with the electromotive force. The electromotive forces corresponding to the two wattless currents, one due to the inductance and the other due to the capacity, are equal and are in opposite phases. We may say that the wattless current required by the inductance is supplied by the capacity (Fig. 312).

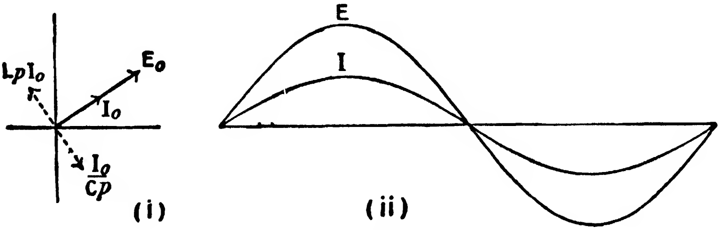


FIG. 312.

Choking Coil.—For many purposes it is required to reduce the current in a given circuit, with a minimum waste of energy, when the current is derived at constant voltage from a given supply. In charging a secondary battery from electric mains, an adjustable resistance is included in the circuit, whose function it is to reduce the current to the required amount, or in other words, to reduce the difference of potential between the ends of the battery to that required for charging. Similarly in running an arc lamp

on a continuous current supply, the arc requires about 40 volts, so that if a current of 10 amperes is to be taken from 100-volt mains, the resistance to be included in the circuit is $\frac{100-40}{10}=6$ ohms. It is merely a matter of applying the ohmic relation $I=\frac{E}{R}$. With an alternating current supply there is another and more economical method which may frequently be employed as an alternative to introducing a resistance; for, let an inductance L be placed in series,

then, $$\frac{I_0}{\sqrt{2}}=\frac{E_0}{\sqrt{2}\sqrt{L^2p^2+R^2}},$$ and in the above case,

$$10=\frac{100}{\sqrt{L^2p^2+4^2}},$$

4 being the effective resistance, $\frac{40}{10}$, of the arc.

$$\therefore L^2p^2+16=100,$$

$$L^2p^2=84, \text{ and } Lp=9{\cdot}2 \text{ approx.}$$

If now the supply has a frequency of 50 cycles per second—

$$p=2\pi \,.\, 50.$$

$$\therefore L=\frac{9{\cdot}2}{2\pi \,.\, 50}=0{\cdot}029 \text{ henry.}$$

Such an inductance is called a choking coil. Its chief advantage lies in the fact that with it, the only waste of energy is due to the hysteresis loss in the iron core (p. 282), which is generally much less than the waste of energy in the resistance that would reduce the current to the same extent as the choking coil. The resistance of the choking coil is generally negligible. When a resistance is used to reduce the voltage, there is a waste of energy I^2R, which in the above example amounts to 600 watts, but with the choking coil the only additional electromotive force introduced differs by 90° in phase with the current, and the effect is therefore wattless.

The virtual volts between the points A and C due to the supply being E volts, that between the ends AB (Fig. 313 (i)) of the non-inductive resistance is RI, and that between the ends BC of the inductance is LpI (p. 345), and the latter differs in phase by 90° from the potential difference RI due to the resistance, since this is in phase with the current. The three electromotive forces are therefore related as the three sides of a right-angled triangle, as shown in Fig. 313 (ii). Thus the sum of the potential differences between A and B, and B and C, is always greater than the

potential difference between A and C. Then we may find Lp graphically by constructing a semicircle upon AC (Fig. 313 (ii)), and making AB equal to the fraction of AC that the required potential difference between the ends of the resistance is of the whole potential difference. On joining CB and dividing its length to scale by I, we obtain Lp.

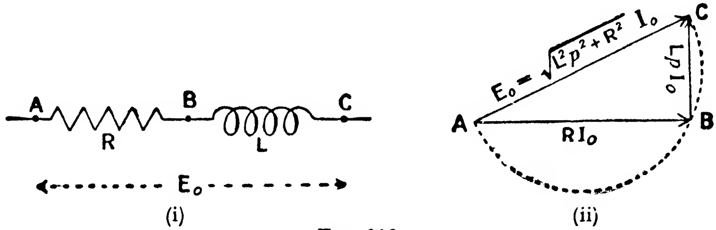


FIG. 313.

Duddell Oscillograph.—Several devices have been employed to determine the wave form of an alternating electromotive force or current, but one of the earliest is that used by W. Duddell[1] in the instrument known as the oscillograph. This is essentially a damped galvanometer, modified to have an exceedingly high frequency of vibration (8000 to 10,000) of the moving part, so that its movement copies the comparatively low frequency couple due to the alternating current.

The phosphor-bronze strip *ssss* passes over the pulley P (Fig. 314), the ends being attached to terminals fixed in the block K. A spring, not shown in the figure, pulls P upwards, and maintains a considerable tension in the strip, whose lower portions are situated in the magnetic field due to a powerful electro-magnet.

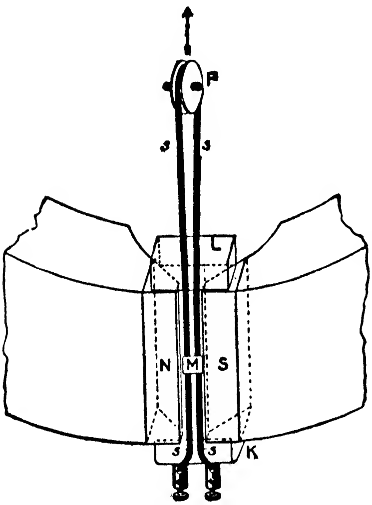


FIG. 314.

On passing a current through the strip, one limb is urged outwards and the other inwards, causing the light mirror M, attached to them, to rotate. The deflection of a beam of light reflected from M is thus proportional at every instant to the current flowing in the strip. The spot of light, falling upon a

[1] W. Duddell and E. W. Marchant, *Inst. El. Eng. Journ.*, **28**, p. 1 1899.

screen or photographic plate describes a straight line when an alternating current is passing.

In order to exhibit the variation of the current, the beam of light is also reflected by a mirror which is rotating about an axis at right angles to the axis of rotation of M, so that the spot of light has a motion upon the screen proportional to time, at right angles to that proportional to the current, and hence describes a path similar to the curves of Fig. 315. The motion of the second mirror is produced by a synchronous motor driven by the alternating current under examination, so that the spot moves over the path repeatedly, an advantage in observing it, as the appearance upon the screen is that of a steady curve.

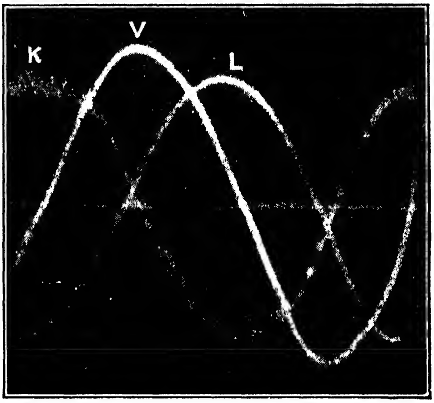


FIG. 315.

For the examination of the electromotive force and current curves simultaneously, two strips such as *s* are placed side by side, the current strip is placed in shunt across a low resistance in which the current to be examined is flowing. The potential strip is placed in series with a high, non-inductive resistance, the ends of this circuit connected to the terminals between which the variation of potential difference to be examined is occurring. The current and potential curves may be arranged to fall simultaneously upon the screen, and in this way the curves in Fig. 315 have been obtained, L being the current curve with large inductance, and K that with large capacity, the lead being nearly 90° for the latter, and the lag nearly 90° for the former (see p. 353). V is the voltage curve.

For the somewhat similar Einthoven string galvanometer see p. 75. Both the Duddell and Einthoven instruments are surpassed in speed of response by the cathode-ray tube (p. 485). The electron beam in this device is usually deflected electrostatically so that negligible current is drawn from the source and its inertia is extremely small. Photographic records are readily obtained.

Transformers.—The industrial use of alternating currents owes its development entirely to the transformer, which is an appliance for converting large current at low voltage to low current at high voltage, and *vice versâ*, with very small loss in energy and without the necessity of moving parts in the appliance. Thus for the

transmission of 10,000 watts at 100 volts the current is 100 amperes, but at 10,000 volts the current is only one ampere. Hence the conductor required in the second case will be much smaller, and therefore less expensive than in the first.

The transformer consists essentially of two coils, a primary and a secondary, wound upon an iron core. The relative number of turns in the two coils depends upon whether it is required to transform up or down in voltage. An ordinary transformer with closed magnetic circuit is shown in Fig. 316, AB being the primary and *ab* the secondary coil.

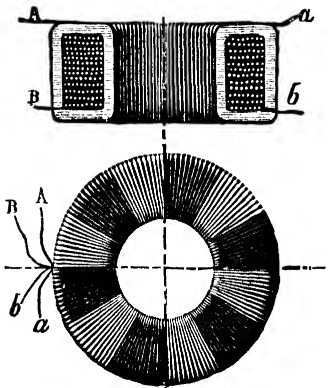


FIG. 316.

We have already found on p. 320 that the electromotive force equations for two circuits having mutual inductance, are—

$$L_1\frac{dI_1}{dt}+M\frac{dI_2}{dt}+R_1I_1=E_1 \quad \text{(i)}$$

$$L_2\frac{dI_2}{dt}+M\frac{dI_1}{dt}+R_2I_2=0\ . \quad \text{(ii)}$$

Let an electromotive force E_1 be applied to the first circuit, there being no source of electromotive force (other than that due to the mutual actions of the currents) applied to the second, and let E_1 vary harmonically. It may be written ${}_0E_1 \sin pt$, and the harmonic solutions for I_1 and I_2 might be obtained as on page 343. But this process is extremely tedious, and the useful results may all be obtained by starting with the currents and afterwards finding the impressed electromotive forces required to produce these currents. Assuming then that the currents I_1 and I_2 differ in phase by the angle θ, we may write them, taking ${}_0I_1$ and ${}_0I_2$ for the maximum values—

$$I_1={}_0I_1 \sin pt \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

and,

$$I_2={}_0I_2 \sin (pt+\theta) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv)}$$

In practice these currents will not be true sine functions, owing to the hysteresis in the iron core of the transformer, but will consist of a sine term together with a higher harmonic due to hysteresis, which is wattless,[1] and may usually be neglected; its study belongs to the province of the electrical engineer.

From equations (iii) and (iv) we get—

$$\frac{dI_1}{dt}=p\,{}_0I_1 \cos pt, \text{ and, } \frac{dI_2}{dt}=p\,{}_0I_2 \cos (pt+\theta).$$

[1] Steinmetz and Berg, "Alternating Current Phenomena."

Substituting in equation (ii) we get—

$$L_2p\,{}_0I_2\cos(pt+\theta)+Mp\,{}_0I_1\cos pt+R_2\,{}_0I_2\sin(pt+\theta)=0,$$

or, $$(L_2p\,{}_0I_2\cos\theta+Mp\,{}_0I_1+R_2\,{}_0I_2\sin\theta)\cos pt$$
$$+(R_2\,{}_0I_2\cos\theta-L_2p\,{}_0I_2\sin\theta)\sin pt=0.$$

This is true for all values of t, and therefore when $pt=\frac{\pi}{2}$, we have—

$$R_2\,{}_0I_2\cos\theta-L_2p\,{}_0I_2\sin\theta=0,$$

or, $$\tan\theta=\frac{R_2}{L_2p},\qquad\text{and }\therefore\ \sin\theta=\frac{R_2}{\sqrt{L_2{}^2p^2+R_2{}^2}}.$$

When $t=0$, we have—

$$L_2p\,{}_0I_2\cos\theta+Mp\,{}_0I_1+R_2\,{}_0I_2\sin\theta=0.$$

And, $$\frac{{}_0I_2}{{}_0I_1}=-\frac{Mp}{L_2p\cos\theta+R_2\sin\theta},$$
$$=-\frac{Mp}{L_2p\cdot\frac{L_2p}{\sqrt{L_2{}^2p^2+R_2{}^2}}+R_2\frac{R_2}{\sqrt{L_2{}^2p^2+R_2{}^2}}},$$
$$\therefore\ {}_0I_2=-\frac{Mp\,{}_0I_1}{\sqrt{L_2{}^2p^2+R_2{}^2}}.$$

From these two results we see that there is a difference in phase between the primary and secondary currents determined by the last relation, together with the fact that $\tan\theta=\frac{R_2}{L_2p}$.

Returning to equation (i), and substituting in it the values of $\frac{dI_1}{dt}$, $\frac{dI_2}{dt}$ and I_1, we have—

$$L_1p\,{}_0I_1\cos pt+Mp\,{}_0I_2\cos(pt+\theta)+R_1\,{}_0I_1\sin pt=E_1,$$
$$\therefore\ E_1=(R_1\,{}_0I_1-Mp\,{}_0I_2\sin\theta)\sin pt+(L_1p\,{}_0I_1+Mp\,{}_0I_2\cos\theta)\cos pt,$$
$$=\left(R_1\,{}_0I_1+Mp\cdot\frac{Mp\cdot{}_0I_1}{\sqrt{L_2{}^2p^2+R_2{}^2}}\cdot\frac{R_2}{\sqrt{L_2{}^2p^2+R_2{}^2}}\right)\sin pt$$
$$+\left(L_1p\,{}_0I_1-Mp\frac{Mp\cdot{}_0I_1}{\sqrt{L_2{}^2p^2+R_2{}^2}}\cdot\frac{L_2p}{\sqrt{L_2{}^2p^2+R_2{}^2}}\right)\cos pt,$$
$$={}_0I_1\left(R_1+\frac{M^2p^2R_2}{L_2{}^2p^2+R_2{}^2}\right)\sin pt+{}_0I_1\left(L_1p-\frac{M^2p^2\,.\,L_2p}{L_2{}^2p^2+R^2}\right)\cos pt,$$
$$={}_0I_1\sqrt{\left(R_1+\frac{M^2p^2R_2}{L_2{}^2p^2+R_2{}^2}\right)^2+p^2\left(L_1-\frac{M^2p^2\,.\,L_2}{L_2{}^2p^2+R_2{}^2}\right)^2}\sin(pt+\phi),$$

where, $$\tan\phi=\frac{\left(L_1-\frac{M^2p^2L_2}{L_2{}^2p^2+R_2{}^2}\right)p}{R_1+\frac{M^2p^2R_2}{L_2{}^2p^2+R_2{}^2}}.$$

We therefore see that the primary electromotive force leads the primary current by an angle ϕ, or, what is the same thing the current lags behind the voltage by this angle. Further, the primary resistance is increased by an effective amount $\frac{M^2p^2R_2}{L_2^2p^2+R_2^2}$ due to the current in the secondary, and the primary inductance is effectively reduced by the amount $\frac{M^2p^2L_2}{L_2^2p^2+R_2^2}$, a phenomenon first pointed out by Maxwell.[1]

Writing $\frac{Mp}{\sqrt{L_2^2p^2+R_2^2}}=P$, we may then express our currents in terms of E_1, taking $E_1={}_0E_1 \sin pt$, and rewriting our equations, and remembering that the current is ϕ in phase behind the electromotive force, (iii) becomes—

$$I_1=\frac{{}_0E_1}{\sqrt{(L_1-P^2L_2)^2p^2+(R_1+P^2R_2)^2}}\sin(pt-\phi) \quad . \quad \text{(v)}$$

$$={}_0I_1 \sin(pt-\phi) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(vi)}$$

where, $\tan\phi=\frac{(L_1-P^2L_2)p}{(R_1+P^2R_2)}$.

Again, since ${}_0I_2=-P\,{}_0I_1$, (iv) becomes—

$$I_2={}_0I_2 \sin(pt-\phi+\theta)$$
$$=-P\,{}_0I_1 \sin(pt-\phi+\theta)$$
$$=P\,{}_0I_1 \sin(pt-\phi+\theta-\pi).$$

For the difference of phase between I_1 and I_2 is the advance θ, of the latter ahead of the former, together with a lag of π implied by the relation ${}_0I_2=-P\,{}_0I_1$.

Now $\tan\theta=\frac{R_2}{L_2p}$, so that $\tan\theta'=\frac{L_2p}{R_2}$, where $\theta=\frac{\pi}{2}-\theta'$, and substituting this value for θ we have—

$$I_2=P\,{}_0I_1 \sin\left(pt-\phi-\theta'-\frac{\pi}{2}\right),$$

and the actual lag of the secondary current behind the primary is an angle $\theta'+\frac{\pi}{2}=\pi-\theta$. The complete expression for I_2 is now—

$$I_2=\frac{{}_0E_1P}{\sqrt{(L_1-P^2L_2)^2p^2+(R_1+P^2R_2)^2}}\sin\left(pt-\phi-\theta'-\frac{\pi}{2}\right) \quad . \quad \text{(vii)}$$

$$={}_0I_2 \sin\left(pt-\phi-\theta'-\frac{\pi}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(viii)}$$

[1] J. C. Maxwell, *Phil. Trans. Roy. Soc.*, **155**, p. 459. 1865.

The meaning of these equations can be more clearly seen by drawing a vector diagram for the electromotive forces in the two circuits (Fig. 317). The equations of electromotive force are (i) and (ii), and the various terms in them and their relative phases may be found from equations (v), (vi), (vii) and (viii). Thus, from (vi) we have

$$R_1 I_1 = R_1\,{}_0I_1 \sin(pt-\phi).$$

Hence we will begin our diagram with the vector OE_0, whose value is ${}_0E_1$, the maximum of E_1. At an angle ϕ behind this we have OA, equal to $R_1\,{}_0I_1$, which is in phase with the primary current, ϕ being very nearly 90° since $L_1 p$ is usually large in comparison with R_1.

Again, from (vi) we have—

$$L_1\frac{dI_1}{dt} = L_1 p\,{}_0I_1 \sin\left(pt-\phi+\frac{\pi}{2}\right) \text{ (see p. 345).}$$

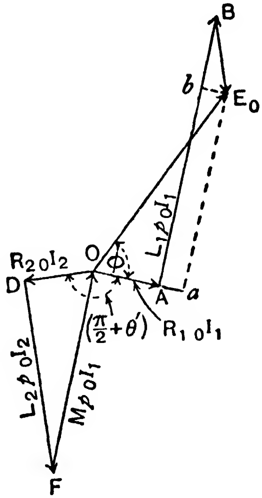


FIG. 317.

Hence AB is drawn 90° ahead of OA, and is made equal to $L_1 p\,{}_0I_1$. From (viii)—

$$R_2 I_2 = R_2\,{}_0I_2 \sin\left(pt-\phi-\theta'-\frac{\pi}{2}\right),$$

so that we next make OD equal to $R_2\,{}_0I_2$ at an angle AOD$=\theta'+\frac{\pi}{2}$ behind OA, and this is in phase with the secondary current. Also from (viii)—

$$L_2\frac{dI_2}{dt} = L_2 p\,{}_0I_2 \sin(pt-\phi-\theta'),$$

and therefore DF is drawn 90° ahead of OD and equal to $L_2 p\,{}_0I_2$. OF is now the vector sum of the two electromotive forces $L_2\frac{dI_2}{dt}$ and $R_2 I_2$ in the secondary, and is equal to the electromotive force due to the variation of current in the primary, which is $M\frac{dI_1}{dt} = Mp\,{}_0I_1 \sin\left(pt-\phi+\frac{\pi}{2}\right)$ from (vi), which we see is parallel to AB; and further the vector BE_0, which is the electromotive force in the primary due to the variation of the secondary current, is $Mp\,{}_0I_2$, and is parallel to DF, since from (viii)—

$$M\frac{dI_2}{dt} = Mp\,.\,{}_0I_2 \sin(pt-\phi-\theta').$$

The three vectors OD, DF and FO have zero resultant, and if

the diagram rotates with constant angular velocity p, their projections on a fixed axis are at any instant the three terms in equation (ii).

Similarly the three vectors OA, AB and BE_0, having a resultant equal to ${}_0E_1$, correspond to the three terms in equation (i).

If a perpendicular E_0a be dropped from E_0 on to OA, the vector Oa is equivalent to $R_1' \, {}_0I_1$ where R_1' is the effective resistance of the primary circuit, which we saw on p. 360 to be $R_1 + \frac{M^2p^2R_2}{L_2^2p^2+R_2^2}$. The vector Aa is therefore equal to $\frac{M^2p^2R_2 \, {}_0I_1}{L_2^2p^2+R_2^2}$.

Similarly E_0a is the quantity $L_1'p \, {}_0I_1$ where L_1' is the effective inductance of the primary, and the vector Bb is therefore equal to $\frac{M^2p^2L_2 \, {}_0I_1}{L_2^2p^2+R_2^2}$, so that the lag ($\phi$) of primary current behind the electromotive force which occurs when the current in the secondary is ${}_0I_2$, might be reproduced when ${}_0I_2$ is zero by increasing the primary resistance by the amount $Aa/{}_0I_1$, and diminishing the primary inductance by the amount $Bb/{}_0I_1$.

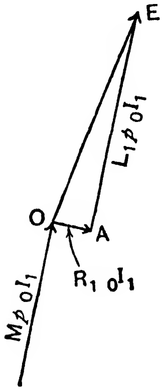


FIG. 318.

When the secondary circuit is broken so that ${}_0I_2$ is zero, the current in the primary is $\frac{{}_0E_1}{\sqrt{L_1^2p^2+R_1^2}}$, and the rate of working is ${}_0E_1 \, {}_0I_1 \cos \phi$, where ϕ is now $\tan^{-1}\frac{L_1p}{R_1}$. The only work done, apart from that due to hysteresis and eddy currents, is that required to drive the current through the primary in opposition to its resistance. The vector diagram reduces to the form shown in Fig. 318. Since R_1 is usually small in comparison with L_1, ϕ is nearly 90°, and becomes more and more nearly equal to 90° as p increases. There is now only one electromotive force in the secondary, equal to $Mp \, {}_0I_1$; and since this is parallel to AE, the electromotive force in the secondary is very nearly the same in phase as the primary electromotive force, OE. From its value $Mp \, {}_0I_1$, we see, calling it ${}_0E_2$, that—

$$ {}_0E_2 = Mp \, {}_0I_1 = Mp \frac{{}_0E_1}{\sqrt{L_1^2p^2+R_1^2}}, $$

and neglecting R_1 in comparison with L_1p—

$$ \frac{{}_0E_2}{{}_0E_1} = \frac{M}{L_1}. $$

When the primary and secondary coils are wound upon the core in such a way that they both enclose the whole of the magnetic flux, we have—

$$L_1L_2=M^2$$

for, $M=\frac{n_2}{n_1}L_1=\frac{n_1}{n_2}L_2$, where n_1 and n_2 are the respective numbers of turns in the primary and secondary, $\therefore\ \frac{{}_0E_2}{{}_0E_1}=\frac{\sqrt{L_1L_2}}{L_1}=\sqrt{\frac{L_2}{L_1}}$.

But the inductance of a coil is proportional to the square of the number of turns linked with the magnetic flux, and therefore $L_1\propto n_1^2$, and $L_2\propto n_2^2$.

$$\therefore\ \frac{{}_0E_2}{{}_0E_1}=\frac{n_2}{n_1}.$$

That is, the electromotive forces in secondary and primary are proportional to their respective numbers of turns.

These conditions may be approximately fulfilled in the case of the induction coil (p. 323), the primary coil having very low and the secondary very high resistance, the latter being very great, not only on account of the number of turns being very great, and the wire employed being thin, but also by the fact that part of the circuit consists of air or some other medium of exceedingly high resistance.

Returning to the vector diagrams (Figs. 317 and 318) we may now consider the effect of reducing the resistance of the secondary until appreciable current passes in it. When this current is small, the vector OD is nearly at right angles to OF, so that the secondary current is nearly opposite in phase to the primary. As work is now being performed in the secondary this implies that more power is taken in at the primary. This is supplied in two ways. In the first place the primary current comes more into phase with the electromotive force, since tan ϕ changes from $\frac{L_1p}{R_1}$ to—

$$\frac{\left(L_1-\frac{M^2p^2L_2}{L_2^2p^2+R_2^2}\right)p}{R_1+\frac{M^2p^2R_2}{L_2^2p^2+R_2^2}},$$

so that ϕ diminishes, and cos ϕ increases, and the power ${}_0E_1\,{}_0I_1\cos\phi$ increases. But in addition to this, the primary current usually increases owing to a diminution in the effective primary impedance from $\sqrt{L_1^2p^2+R_1^2}$ to $\sqrt{(L_1-P^2L_2)^2p^2+(R_1+P^2R_2)^2}$.

If the latter of these two quantities is less than the former—

$$L_1^2p^2+R_1^2>L_1^2p^2-2L_1L_2P^2p^2+P^4L_2^2p^2+R_1^2+2R_1R_2P^2+P^4R_2^2.$$

$$\therefore\ 2L_1L_2p^2>P^2(L_2^2p^2+R_2^2)+2R_1R_2.$$

But,
$$P^2=\frac{M^2p^2}{L_2^2p^2+R_2^2}.$$

$$\therefore\ 2L_1L_2p^2>M^2p^2+2R_1R_2.$$

In the case of an ordinary transformer, R_1 and R_2 are made as small as possible, in order to avoid loss of energy due to heating of the conductors. Therefore, neglecting the term $2R_1R_2$, we have—

$$2L_1L_2>M^2,$$

a condition which is necessarily fulfilled since M^2 cannot be greater than L_1L_2. Thus the result of the current in the secondary is to decrease the effective impedance of the primary, and the primary current therefore increases. It should be noted that if the resistances are not small it does not follow that the effective impedance of the primary is reduced and the current increased. There is, however, always an increase in the power absorbed by the primary, on account of the advance in phase of the primary current caused by the current in the secondary.

For an account of the efficiency of transformers and the measurement of the various losses occurring in them the student is referred to works on electrical engineering.

Resistance and Inductance of Wires for Currents of High Frequency.—A steady current flowing in a uniform wire is distributed uniformly in the cross-section of the wire, the current density being constant over any given section. When the voltage applied between the ends of the wire is alternating, the distribution of current is no longer uniform ; there is a concentration of the current in the outer layers, and, when the frequency is very great, the current is almost entirely confined to the surface layer. This phenomenon is known as the "*skin effect,*" and on account of it, the effective resistance of the wire is greatly increased. For this reason, conductors that are required to carry high frequency alternating currents are made up of a number of strands of fine wire, insulated from each other, in order to have a large surface for any given area of cross-section, since the central parts of thick wires would not carry any appreciable part of the current and would therefore be useless.

The reason for this distribution of the current may be understood by examining Fig. 219, in which the magnetic field for a wire carrying steady current is shown. Taking any cylindrical shell of the wire, the magnetic field outside it is the same as

though the current in the shell were all at the axis of the wire, but for points inside the shell the field is zero. Thus for a given current, the total magnetic flux is greater when the current flows along the axis than when it flows in a surface layer of the wire, by the amount of flux that fills the space occupied by the wire when the current flows along the axis.

If then we imagine two elements of the wire of equal cross-section, one constituting the central portion A, in Fig. 319, and the other a cylindrical shell B, these will carry equal currents when the electromotive force is steady, but A will have a greater self-inductance than B. When the electromotive force is alternating, B will therefore have the less impedance, and more current will flow in it; and further, the current in A will lag behind that in B, owing to the greater inductance of A. The phase of the current gets later and later as we pass from the surface of the wire to the interior. This is entirely in accordance with the fact that the flow of energy from the source of electromotive force takes place in the dielectric surrounding the wire, time being necessary for it to penetrate from the surface to the interior. This aspect of the question will be studied later (see p. 426).

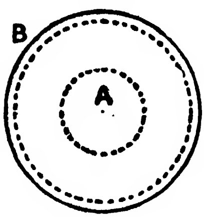


FIG. 319.

If we consider a conductor consisting of the two parts A and B only, where the resistance of each for steady current is R, the inductance of A being L, that of B is equal to M, the mutual inductance of the two parts; since the magnetic flux linked with the circuit whose cross-section is B, is also linked with A, and if we apply an alternating electromotive force $E_0 \sin pt$ to the two in parallel, we see by the equations on p. 360 that the effective resistance of A is—

$$R+\frac{M^2p^2R}{M^2p^2+R^2}=R\left(1+\frac{M^2p^2}{M^2p^2+R^2}\right),$$

and its effective inductance is—

$$L-\frac{M^2p^2M}{M^2p^2+R^2},$$

while the effective resistance and inductance of B are respectively—

$$R+\frac{M^2p^2R}{L^2p^2+R^2}=R\left(1+\frac{M^2p^2}{L^2p^2+R^2}\right),$$

and,

$$M-\frac{M^2p^2L}{L^2p^2+R^2}.$$

The total resistance is therefore in each case increased, and the

effective inductance reduced, the change depending on the square of p; but since $L>M$, the resistance of A rises more rapidly than that of B when p increases, so that the current becomes at very high frequency almost entirely confined to B. The better the conductivity of the material the smaller is the value of R, and therefore the greater is the quantity $\frac{M^2p^2M}{M^2p^2+R^2}$ or $\frac{M^2p^2L}{L^2p^2+R^2}$, the limiting values when $R=0$ being M and $\frac{M^2}{L}$. Since $\frac{M}{L}<1$, M is greater than $\frac{M^2}{L}$, and the concentration of the current into B is greater when R is small than when it is great, for in this latter case $\frac{M^2p^2M}{M^2p^2+R^2}$ and $\frac{M^2p^2L}{L^2p^2+R^2}$ approach to equality.

If the wire consist of a material of high permeability, L is enormously increased while M is practically unchanged. This again accentuates the crowding of the current into the outer layers; in other words it increases the skin effect.

The problem of the distribution of an alternating current in an actual wire is beyond the scope of this book, but it may be pointed out that in the case of a straight circular wire carrying alternating current, the effective resistance R′ may be calculated from a relation which may be written in the form—

$$R'=R\left\{1+\frac{1}{12}\left(\frac{2\pi^2 n\mu a^2}{\rho}\right)^2-\frac{1}{180}\left(\frac{2\pi^2 n\mu a^2}{\rho}\right)^4+\ \ldots\ \right\},$$

given by the late Lord Rayleigh,[1] where R is the resistance for steady current, a being the radius of the wire in centimetres, n the frequency of alternation, and ρ the specific resistance of the material of the wire in absolute units, provided that $\frac{4\pi^2a^2\mu n}{\rho}$ is not greater than 5. For very high frequencies the relation is—

$$R'=R\sqrt{\frac{\pi^2 n\mu a^2}{\rho}}.$$

The effective resistance of a straight wire for high frequency currents has been measured by Sir Ambrose Fleming [2] by comparing the currents in two similar wires that produce heating at the same rate, the current in one wire being steady and the other oscillating. The arrangement employed is shown diagrammatically in Fig. 320. The two wires AB and CD are situated in glass tubes which are united by a bent tube containing a little paraffin oil, with an air bubble at P. The currents in AB and

[1] Lord Rayleigh, *Phil. Mag.* (Ser. 5), **21**, p. 381. 1886.
[2] J. A. Fleming, *Proc. Phys. Soc. Lond.*, **23**, p. 103. 1911.

CD are adjusted until the air bubble remains in its equilibrium position, the system then acting as a differential air thermometer; it indicates that the rates of production of heat in the wires are equal.

The alternating current is measured by an ammeter of the type described on p. 222, which measures virtual amperes. It consists in the oscillatory discharge from a condenser.

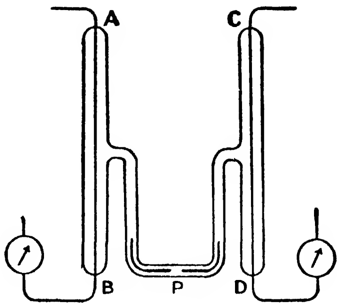


FIG. 320.

For equal heating in the two wires, $I^2R=I'^2R'$ where I' is the virtual current, and R' the effective resistance, I and R being the values for steady current.

Since the wires and the vessels may not be identical in the rates at which the heat produced in the wires is dispersed, the currents are interchanged, so that the wire which previously carried the alternating current now carries the steady current, the effect of dissimilarity in the tubes and wires being in this way eliminated. The observations were in very good agreement with the calculated results. In the case of a bare copper wire of diameter 0·03149 cm. the resistance for a frequency of $1{\cdot}08\times10^6$ is 1·45 times that for steady current, while for a diameter of 0·198 the ratio is 8·10.

Shielding Effect of a Mass of Metal.—The presence of a mass of conducting material in the neighbourhood of a circuit carrying alternating current produces an effect which may be understood from the equations on p. 360. The effective self-inductance of the circuit is reduced, for the induced current in the material has a magnetic field which is opposite in sign to that due to the current in the circuit, and consequently while the current is growing, the magnetic flux linked with the circuit is less than would be the case if the mass of metal were absent. The back electromotive force, due to the growth of the magnetic flux, is therefore reduced, which is equivalent to saying that the effective self-inductance is diminished. Similarly at stopping, the induced current is in the direction of that in the circuit (see Fig. 297), and the flux dies away at a lessened rate.

As the conductivity of the material increases, so the effect is enhanced, for the induced currents become greater, while the reverse is the case when the conductivity is diminished. For this reason it is necessary to avoid using continuous masses of metal in the construction of coils of large inductance. The frame on which the coil is wound should be of some non-

conducting material, and if this is inconvenient, a saw cut should be made in a direction at right angles to the direction of the electromotive force produced by the varying magnetic flux. This enormously diminishes the effect. Further, if an iron core is employed, it should be laminated or else built up of wires insulated from each other. A slight film of oxide on the wires or laminæ will be highly beneficial. For this reason the cores of transformers are usually laminated, for not only is the effect of these eddy currents upon the inductance objectionable, but they involve a waste of energy, the current in the metal involving a conversion of electrical energy into heat within the material.

In a similar manner a sheet or mass of highly conducting metal may be used to screen a given space from the effects of an alternating magnetic field.

Let an alternating current $I_1 = {}_0I_1 \sin pt$, be flowing in a given circuit; then for a neighbouring circuit or mass of metal, the equation of electromotive forces will be—

$$L_2\frac{dI_2}{dt}+M\frac{dI_1}{dt}+R_2I_2=0.$$

When R_2 is very small, as we may suppose it to be in the given case—

$$L_2\frac{dI_2}{dt}+M\frac{dI_1}{dt}=0.$$

$$\therefore\ L_2I_2+MI_1=\text{constant}.$$

L_2I_2 is the magnetic flux through the closed circuit due to the current I_2 in it, and MI_1 is the flux due to the current I_1, and since the sum of these two is constant, the variation in flux, which, without the closed circuit would be

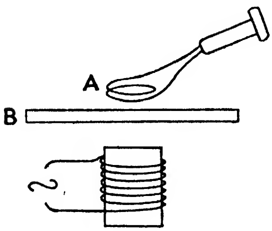


FIG. 321.

$$M_0I_1 \sin pt,$$

is reduced to zero. This effect may be demonstrated by placing a coil A (Fig. 321) which is in series with a telephone receiver, near an electromagnet excited by an alternating current. A note whose pitch is equal to twice the frequency of alternation of the current will be heard in the telephone. If now a thick sheet of copper, B, be placed near A, the loudness of the note is much reduced. The sound will never disappear, for the sheet cannot have absolutely zero resistance, and consequently the limiting condition implied by the last equation is never reached.

It may be noted that unless the alternations are extremely rapid, the effect of the presence of the plate is considerable, on

whichever side of the coil A it is placed, and is greater the nearer it is to A. If the magnet is at some distance, it is immaterial on which side of A the plate is situated, for it is the variation in flux at B that is reduced, and there is a consequent reduction in the variation at all points near B. A radial slot cut in the plate B enormously increases its resistance to the induced currents and correspondingly diminishes the effect.

Repulsion between Conductor and a Circuit carrying Alternating Current.—Closely allied to the last described effect is the phenomenon of repulsion that occurs between a circuit carrying an alternating current, and a conductor. A closed circuit or conductor A (Fig. 322) situated near an alternating electromagnet, is threaded by a magnetic flux $MI_0 \sin pt$, due to the current $I_0 \sin pt$, in the electromagnet.

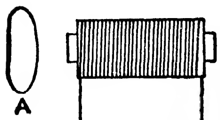


FIG. 322.

Since A has very small inductance, its reaction upon the electromagnet is infinitesimal, and we may consider that an alternating electromotive force $-M \cdot \frac{dI}{dt} = -MpI_0 \cos pt = MpI_0 \sin\left(pt - \frac{\pi}{2}\right)$ acts in it. This electromotive force is 90° in phase behind the current I_0, and neglecting A's inductance, the current in it will be in phase with this electromotive force. Thus—

$$I_2 \propto MpI_0 \sin\left(pt - \frac{\pi}{2}\right).$$

There will be a force between the two currents, proportional to their product I_1I_2.

$$\therefore \text{ Force} \propto MpI_0^2 \sin pt \sin\left(pt - \frac{\pi}{2}\right).$$

We have already seen (p. 349) that the mean value of a quantity such as $\sin pt \sin\left(pt - \frac{\pi}{2}\right)$, in which the two harmonic components differ in phase by 90°, is zero, and therefore if the inductance of A be zero, the mean force on it is also zero. But the inductance, although small, cannot be zero, and the current I_2 therefore lags in phase by more than 90° behind I_1:

$$\therefore \text{ Force} \propto MpI_0^2 \sin pt \sin\left(pt - \frac{\pi}{2} - \theta\right),$$

$$\propto MpI_0^2 \sin^2 pt \cos\left(\frac{\pi}{2} + \theta\right) - \frac{MpI_0^2}{2} \sin 2pt \sin\left(\frac{\pi}{2} + \theta\right).$$

The mean value of the last term we saw on p. 346 to be zero, and the mean of the first is $\frac{1}{2}MpI_0^2 \cos\left(\frac{\pi}{2} + \theta\right)$. Since $\cos\left(\frac{\pi}{2} + \theta\right)$

is necessarily negative when θ is small, the force on A is a repulsion, and the coil A experiences in each complete cycle, an impulse pushing it away from the magnet. The curves for I_1 and I_2 are drawn in Fig. 323, and the dotted curve is drawn for the product $I_1 I_2$. Owing to the phase difference between I_1 and I_2 being greater than 90° (compare with Fig. 309, where the phase difference is less than 90°) the positive loops *a*, which indicate an attraction owing to the currents being in the same direction, are smaller than the negative loops *b* which

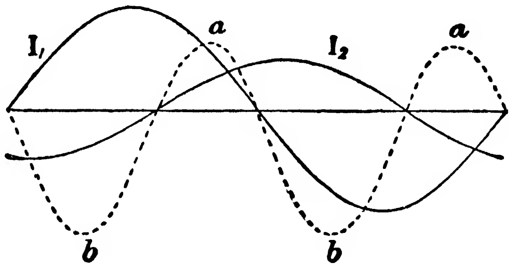


FIG. 323.

indicate a repulsion. It is thus seen that the repulsions predominate.

It may be seen from the expression for the force, that this increases with p, and with the square of the current I_0. Hence, to get large effects, currents of considerable value and of high frequency are necessary.

The resistance of A plays an important part in the phenomenon, for the lower the resistance the greater the induced current in it and the larger the effect. But the lower the resistance, the greater is the lag of the current behind the electromotive force owing to the increase in the time constant, and hence the lag θ increases, the result being that the loops *b* in Fig. 323 are increased while the loops *a* are diminished, the repulsion being still further increased.

By means of powerful alternating-current electromagnets, metal rings of considerable weight may be supported.

Rotating Magnetic Field.—If two coils carrying alternating currents are placed at right angles to each other, the resulting magnetic field at any instant may be found by compounding the fields due to the two coils, according to the ordinary law of addition of vector quantities or parallelogram of forces. When the currents have the same frequency, the resultant magnetic field at any point is periodic, and has the same frequency as the currents.

Representing the currents by the equations $I_1 = {}_0I_1 \sin (pt + \theta)$

and $I_2={}_0I_2 \sin pt$, the magnetic fields are in phase with the respective currents and may be represented by—

$$H_1={}_0H_1 \sin (pt+\theta)$$

and,
$$H_2={}_0H_2 \sin pt.$$

If the field H_1 is due to the current in the coil AB (Fig. 324) it may be represented by the vector OH_1, and the field due to the coil CD is represented by OH_2, and at the instant to which the diagram refers, OH' is the resultant field. This has the value $\sqrt{H_1^2+H_2^2}$, and is inclined to the direction of OH_1 by an angle ϕ whose tangent is $\frac{H_2}{H_1}$, that is, $\tan \phi=\frac{H_2}{H_1}$. Both H' and ϕ, which define the resultant field, vary periodically; for when $t=0$, then $H_2=0$, and

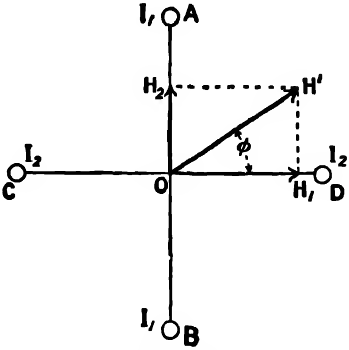


FIG. 324.

$$H'={}_0H_1 \sin \theta.$$

Also when $t=-\frac{\theta}{p}$, $H_1=0$, and $H'={}_0H_2 \sin (-\theta)$.

Thus the resultant field H' rotates, and at the same time varies in magnitude.

The case of most importance is that in which ${}_0H_1={}_0H_2=H_0$.

Then,
$$H'=H_0\sqrt{\sin^2 pt+\sin^2(pt+\theta)},$$

and,
$$\tan \phi=\frac{\sin pt}{\sin (pt+\theta)};$$

and if in addition $\theta=\frac{\pi}{2}$, then—

$$H'=H_0, \text{ and, } \tan \phi=\frac{\sin pt}{\cos pt}=\tan pt,$$

that is, $\phi=pt$. The resultant field is therefore constant in value, and rotates with constant angular velocity p.

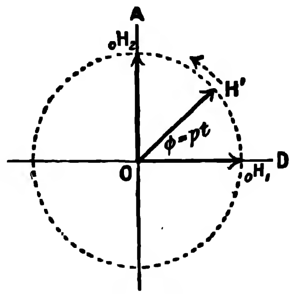


FIG. 325.

Then at time $t=0$, $H_1=H_0$, and $H_2=0$, and the position of H' is O_0H_1 (Fig. 325). At time $t=\frac{\pi}{2p}$, $H_1=0$, and $H_2=H_0$, and the position of H' is O_0H_2. Hence the direction of

rotation of H′ is positive, that is from OD to OA, or anti-clockwise, when H_1 is 90° in phase ahead of H_2.

If, on the other hand, H_1 lags 90° in phase behind H_2—

$$H_1 = H_0 \sin\left(pt - \frac{\pi}{2}\right) = -H_0 \cos pt,$$

and,
$$H_2 = H_0 \sin pt,$$
$$\tan\phi = -\tan pt, \text{ and } \phi = -pt.$$

The direction of rotation of the resultant field H′ is therefore in this case negative, that is, opposite to that of the rotation of the vectors ${}_0H_1$ and ${}_0H_2$, and its angular velocity is $-p$.

A conductor placed at O will consequently experience a couple, which is in all cases in the direction of rotation of the resultant field, for an exactly similar reason to that for which the magnet rotates in Arago's experiment, described on p. 253. In this case the field rotates, whereas in Arago's experiment the conductor rotates, but in both experiments it is the relative motion that determines the couple between them, for by Lenz's law (p. 253) we know that the forces due to the magnetic effects between the field due to the induced currents and the original field, are such as to oppose the relative motion of the field and conductor.

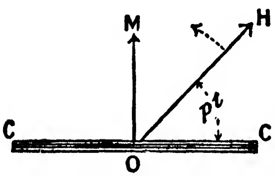


Fig. 326.

The couple acting on a mass of metal situated in a rotating magnetic field cannot in general be calculated, because of the difficulty in finding the distribution of the induced currents in the body, but if we take a plane coil CC (Fig. 326) the problem becomes much simpler.

The magnetic flux passing through the coil at any instant is AH sin pt, where pt is the angle between H and the plane of the coil, and A the area of the coil. The electromotive force in the coil is given by the equation—

$$e = -\frac{dN}{dt},$$

$$\therefore e = -\frac{d}{dt}(AH \sin pt)$$

$$= -AHp \cos pt = AHp \sin\left(pt - \frac{\pi}{2}\right).$$

The current in the coil is therefore—

$$i = \frac{AHp}{\sqrt{l^2p^2 + r^2}} \sin\left(pt - \frac{\pi}{2} - a\right) \quad . \quad . \quad . \quad \text{(see p. 344)},$$

where l and r are the inductance and resistance of the coil and $\tan \alpha = \frac{lp}{r}$.

The magnetic moment of the coil is Ai (see p. 225), and is perpendicular to its plane, being directed along the normal in direction MO. The couple acting on the coil, tending to turn it into the direction of H, is then—

$$-\mathrm{A}i\mathrm{H}\cos pt = -\frac{\mathrm{A}^2\mathrm{H}^2 p}{\sqrt{l^2p^2+r^2}}\sin\left(pt-\frac{\pi}{2}-\alpha\right)\cos pt$$

$$= -\frac{\mathrm{A}^2\mathrm{H}^2 p}{\sqrt{l^2p^2+r^2}}\left\{\sin pt\cos pt\cos\left(\frac{\pi}{2}+\alpha\right)-\cos^2 pt\sin\left(\frac{\pi}{2}+\alpha\right)\right\}.$$

We have seen that the mean value of $\sin pt \cos pt$, or $\sin 2pt$, for a cycle is zero, while that of $\sin^2 pt$ or $\cos^2 pt$ is $\frac{1}{2}$.

Therefore mean couple c is given by—

$$c = \frac{\mathrm{A}^2\mathrm{H}^2 p}{2\sqrt{l^2p^2+r^2}}\sin\left(\frac{\pi}{2}+\alpha\right) = \frac{\mathrm{A}^2\mathrm{H}^2 p}{2\sqrt{l^2p^2+r^2}}\cos\alpha.$$

But, $$\tan\alpha = \frac{lp}{r},$$

$$\therefore \cos\alpha = \frac{r}{\sqrt{l^2p^2+r^2}},$$

and, $$c = \frac{\mathrm{A}^2\mathrm{H}^2 rp}{2(l^2p^2+r^2)}.$$

The mean couple is in the direction of rotation of the field.

The average couple therefore depends on the value of p, and this is the relative angular velocity of the field with respect to the coil. It is evidently zero when $p=0$, and again when p is infinite. If then the coil is mounted so that it can rotate, its angular velocity in the direction of rotation of the field will increase until the rate at which work is done in opposition to friction of all kinds is equal to that done by the rotating field. On releasing the coil its angular velocity will increase at first, and as a result p, the relative angular velocity of the field with respect to the coil, diminishes and the couple still further increases; but the speed will never be equal to that of the field since in this case p would be zero and the couple would vanish. The average couple is a maximum when p has some value between zero and infinity.

To find the value of p for the average couple to be a maximum, find the condition that the rate of change of the average couple with respect to p shall be zero; *i.e.* let—

$$\frac{dc}{dp} = 0.$$

Now,
$$c=\frac{A^2H^2rp}{2(l^2p^2+r^2)},$$
$$\therefore \frac{dc}{dp}=\frac{A^2H^2r}{2}\frac{d}{dp}\left(\frac{p}{l^2p^2+r^2}\right)$$
$$=A^2H^2r\frac{l^2p^2+r^2-2l^2p^2}{2(l^2p^2+r^2)^2},$$

and putting this equal to zero we have—
$$l^2p^2=r^2$$
and,
$$p=\frac{r}{l}.$$

Again, $\tan\alpha=\frac{lp}{r}$, so that the value of α for maximum couple is $\tan^{-1}=45°$, and the value of the couple under these circumstances is, since $p=\frac{r}{l}$—
$$c=\frac{A^2H^2r\,.\,r}{4r^2l}=\frac{A^2H^2}{4l}.$$

This does not mean that the couple is independent of the resistance, which would obviously be incorrect, but merely gives the value of the couple when the condition $lp=r$ is fulfilled.

That this condition corresponds to a maximum couple may be proved by obtaining $\frac{d^2c}{dp^2}$ and showing that $\frac{dc}{dp}$ is decreasing when $lp=r$.

The quantity $\frac{p}{l^2p^2+r^2}$ has been plotted as ordinate in Fig. 327 for a case in which $R=1$ ohm or $r=10^9$ absolute units and $L=0{\cdot}1$ henry $=10^8$ absolute units. The frequency of alternation, n, has been taken as abscissa, the value of p being $2\pi n$. The relation
$$\frac{r}{l}=p=2\pi n$$
gives the maximum couple at the frequency $n=\frac{10}{2\pi}=1{\cdot}59$, and it will be seen that the quantity has the value 5×10^{-18} at this frequency, and the average couple is $\frac{1}{2}A^2H^2r\times5\times10^{-18}$. Now H may in a practical case be, say, 1000, and if the coil have an effective area of, say, 10,000 sq. cm.—
$$\text{maximum average couple}=\tfrac{1}{2}10^8\times10^6\times10^9\times5\times10^{-18}$$
$$=2{\cdot}5\times10^5 \text{ C.G.S. units.}$$

It should be noticed that the maximum occurs very near the axis

of zero frequency, and from this maximum the couple falls gradually as the angular velocity of the field relatively to the coil increases.

The motion of a conductor in a rotating magnetic field has been put to many uses, the most notable of which is the construction of electro-motors in which two electromagnets are traversed by alternating currents of different phases. A mass of metal or a system of closed coils is mounted upon an axle in the rotating field, and experiences a driving couple as explained above.

If a conductor or coil E be mounted between two pairs of magnets AB and CD (Fig. 328) carrying alternating currents

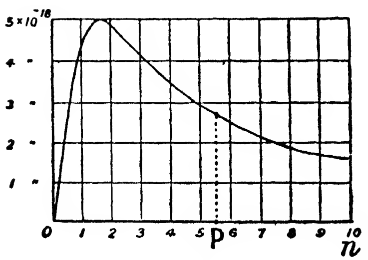


FIG. 327.

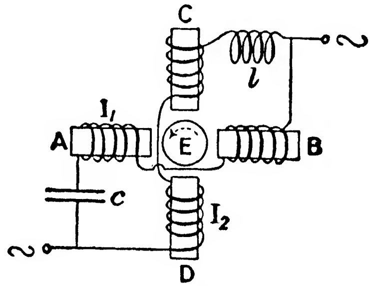


FIG. 328.

differing in phase, the resulting magnetic field at E is a rotating field, and the conductor rotates. The magnets A and B may with advantage be combined to form one field magnet, as also may C and D, but they are represented as separate magnets in the diagram for the sake of clearness.

If the alternating currents in the two circuits are derived from the same supply, the currents in the two pairs of magnets will not differ in phase unless the time constants of the two circuits differ. To make the time constants differ, an extra inductance l may be introduced into the CD circuit, or a capacity into the circuit AB, or both.

Thus, $$I_1={}_0I_1 \sin (pt+\theta_1)$$

and, $$I_2={}_0I_2 \sin (pt-\theta_2).$$

If ${}_0I_1$ and ${}_0I_2$ are equal and the magnets similar, then the magnetic field is a simple rotating field, when—

$$\theta_1+\theta_2=\frac{\pi}{2}.$$

In using motors of this type, the supply usually consists of two separate currents carried by two distinct circuits, the currents

differing in phase by 90°. Such motors having circuits with alternating currents in different phases are called polyphase motors, and have the great convenience that they will start under load.

Single-Phase Motor.—A single alternating magnetic field may be looked upon as the resultant of two equal fields rotating with equal angular velocities in opposite directions. If the two coincide when in the direction Oy (Fig. 329) then, when one of them (OA) makes angle pt with Ox the inclination of the other (OB) to Ox is $\pi - pt$. Thus the components parallel to Ox are $H_0 \cos pt$ and $-H_0 \cos pt$, and these always annul each other; while the components parallel to Oy are $H_0 \sin pt$ and $H_0 \sin (\pi - pt)$, and these added together give the alternating field $2H_0 \sin pt$. A mass of metal or coil of wire mounted in the field so that it can rotate, will experience equal and opposite couples due to the oppositely rotating components of the field, and it will therefore remain at rest.

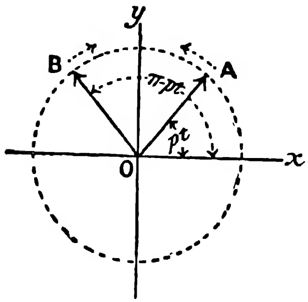


FIG. 329.

Let the angular velocity of the fields with respect to the coil be represented by the value P on the curve (Fig. 327), and it will be seen that if the coil be given an angular velocity p_1 in one direction or the other, the relative angular velocity of one of the components of the field with respect to the coil will become $P - p_1$ and the other $P + p_1$. The couple produced by the former is, therefore, as will be seen from the curve, greater than that for the latter, and the coil will gain velocity. The couple due to the driving component thus increases and the other diminishes until the mechanical work done on account of rotation prevents further increase in velocity.

The coil will therefore run as a motor in the direction in which it is given a start, but owing to the smallness of the resulting couple until the angular velocity approaches that of the driving component of the field, such a motor will not start under load. Hence such motors are provided with a second magnetising coil and some such device as that described on p. 375, for producing a difference of phase between the currents in the two coils. A single rotating magnetic field is therefore created, and the motor starts. When running at sufficient speed, the second or starting circuit is cut out and the motor continues to run as a single-phase machine.

Imaginary Quantities.—The usefulness of the exponential forms of the sine and cosine has already been seen (p. 334), and we will

now make a further application of them to the problems of alternating currents. Let us consider the imaginary quantity $\sqrt{-b^2}$. This has no real value, but may be defined as the quantity whose square is equal to $-b^2$. Writing it in the form $\sqrt{-1}\ b$, or jb, where $j=\sqrt{-1}$, we can see that $j=\sqrt{-1}$, $j^2=-1$, $j^3=-\sqrt{-1}$, $j^4=1$, etc.

If we multiply any vector, say A, by j^2, we obtain $j^2\text{A}=-\text{A}$. Thus the sign is reversed, which is equivalent to a reversal in direction of the vector, or a rotation of its direction through 180°. Multiplying again by j^2 or -1, it again becomes $+\text{A}$, and has therefore been rotated through a further 180°. From this we see that j^2 may be looked upon as an operator, the effect of which is to rotate any vector upon which it operates, through 180°. Similarly, j rotates it through 90°, j^3 through 270°, etc.

Any quantity whatever may be written in the form $a+jb$, where a is a real quantity and jb imaginary, and, further, if any equation involves both real and imaginary quantities, the sum of the real quantities is zero, and likewise that of the imaginaries.

Thus, if $\quad a+jb=a'+jb'$,

then, $\quad a-a'=j(b'-b)$,

which cannot be true unless $a-a'=0$, and $b'-b=0$, for otherwise we should have a real quantity equal to an imaginary.

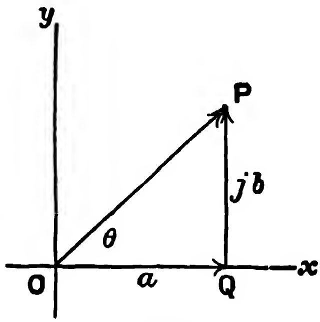


FIG. 330.

If a and b are vectors in the same direction, jb and a are vectors at right angles to each other, and the position of any point P (Fig. 330) may be represented by the vector $a+jb$, since $\text{OQ}=a$ and $\text{QP}=jb$.

$$\therefore \text{ vector OP}=a+jb.$$

$$\text{Again, if OP}=r\text{, and angle QOP}=\theta,$$

$$\text{vector OP}=r(\cos\theta+j\sin\theta).$$

r is usually called the modulus and θ the argument of the complex quantity represented by the vector OP. From Fig. 330 we see that $r^2=a^2+b^2$ and $\tan\theta=\dfrac{b}{a}$, and hence the modulus and argument of any imaginary quantity of the form $a+jb$ are known.

Rotating Vector.—The exponential forms for $\sin\theta$ and $\cos\theta$ are $\dfrac{\epsilon^{j\theta}-\epsilon^{-j\theta}}{2j}$ and $\dfrac{\epsilon^{j\theta}+\epsilon^{-j\theta}}{2}$ respectively, and employing these

forms, the quantity $r(\cos\theta+j\sin\theta)$ becomes $r\epsilon^{j\theta}$, and $r(\cos\theta-j\sin\theta)$ becomes $r\epsilon^{-j\theta}$.

If the vector OP, or r, rotates with angular velocity p, and t is the interval of time since it coincided with Ox, then $\theta=pt$, and $r(\cos pt+j\sin pt)=r\epsilon^{jpt}$. Now, $r\cos pt$ is the projection of r upon the axis of x at any instant, and is a quantity which varies harmonically; it is also the real part of the complex quantity $r\epsilon^{jpt}$. Similarly, $r\sin pt$ is the projection upon the axis Oy. The real part of $r\epsilon^{jpt}$ is therefore a harmonic motion taking place in the direction of the axis of x, and the imaginary part a similar harmonic motion, a quarter of a period later, in the axis of y.

Application of Imaginaries to Circuit having Inductance, Capacity and Resistance.—The alternating electromotive force $E_0\cos pt$ may be looked upon as the real part of the quantity $E_0\epsilon^{jpt}$, or the projection upon the axis of x, of this rotating vector. The equation of electromotive forces (p. 352) may therefore be written—

$$L\frac{dI}{dt}+RI+\frac{Q}{C}=E_0\epsilon^{jpt}. \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

Now consider a solution, $I=A\epsilon^{jpt}$, which has evidently the same periodicity as the electromotive force. We must find the nature of the quantity A.

$$\frac{dI}{dt}=jpA\epsilon^{jpt}, \text{ and, } Q=\int Idt=\frac{A}{jp}\,.\,\epsilon^{jpt}+\text{const.}$$

The constant in the last expression must be zero, since we are dealing entirely with harmonic changes;

$$\therefore\ Q=\frac{A}{jp}\epsilon^{jpt}=-\frac{jA}{p}\epsilon^{jpt}.$$

Equation (i) then becomes, on substitution—

$$LpjA\epsilon^{jpt}+RA\epsilon^{jpt}-\frac{jA}{Cp}\epsilon^{jpt}=E_0\epsilon^{jpt},$$

or,
$$A=\frac{E_0}{\left(Lp-\frac{1}{Cp}\right)j+R},$$

so that
$$I=\frac{E_0}{\left(Lp-\frac{1}{Cp}\right)j+R}\epsilon^{jpt}.$$

$\dfrac{1}{\left(Lp-\frac{1}{Cp}\right)j+R}$ is a complex quantity, and may be written,

$$\frac{R-\left(Lp-\frac{1}{Cp}\right)j}{\left\{\left(Lp-\frac{1}{Cp}\right)j+R\right\}\left\{R-\left(Lp-\frac{1}{Cp}\right)j\right\}}=\frac{R-\left(Lp-\frac{1}{Cp}\right)j}{\left(Lp-\frac{1}{Cp}\right)^2+R^2}$$

$$=a-jb=r\epsilon^{-j\theta}$$

$$\therefore\ r=\sqrt{a^2+b^2}=\sqrt{\frac{R^2+\left(Lp-\frac{1}{Cp}\right)^2}{\left\{\left(Lp-\frac{1}{Cp}\right)^2+R^2\right\}^2}}=\frac{1}{\sqrt{\left(Lp-\frac{1}{Cp}\right)^2+R^2}}.$$

and $$\theta=\tan^{-1}\frac{b}{a}=\tan^{-1}\frac{Lp-\frac{1}{Cp}}{R}$$

$$\therefore\ I=\frac{E_0}{\sqrt{\left(Lp-\frac{1}{Cp}\right)^2+R^2}}\epsilon^{j(pt-\theta)}.$$

The real part of this is the harmonic current required, and this will be seen to be in agreement with the solution on p. 353. The phase difference θ must not be confused with the θ of Fig. 330.

When the electromotive forces and currents in a circuit are entirely simple harmonic, and the inductance, resistance and capacity are in series, the quantity E, which is equivalent to $L\frac{dI}{dt}+RI+\frac{Q}{C}$, is numerically equal to the difference of potential between the ends of the circuit.

$$\text{If } I=I_0\epsilon^{jpt},\text{ then } \frac{dI}{dt}=jpI_0\epsilon^{jpt}=jpI\,;\ Q=\int I_0\epsilon^{jpt}dt=\frac{I_0\epsilon^{jpt}}{jp}=\frac{I}{jp}$$

$$\therefore\ E=LjpI+RI+\frac{I}{Cjp}=\left(Ljp+R+\frac{1}{Cjp}\right)I.$$

If two circuits are in parallel, the difference of potential between the ends is the same for both, at every instant, and the manner in which the current divides between them is given by the relation,

$$\left(L_1jp+R_1+\frac{1}{C_1jp}\right)I_1=\left(L_2jp+R_2+\frac{1}{C_2jp}\right)I_2.$$

It follows that for circuits in which the effects of mutual inductance may be neglected, the alternating current may be calculated from the applied harmonic electromotive force by a law corresponding to Ohm's law (p. 59), where the operator $\left(Ljp+R+\frac{1}{Cjp}\right)$ takes the place of the simple resistance.

General Form of Wheatstone's Bridge.—It was pointed out by Heaviside [1] that the Wheatstone's bridge might be balanced when the arms contain inductances and capacities, provided that one branch is electrically a copy, on some definite scale, of the other branch.

A general form of the Wheatstone's bridge is shown in Fig. 331, in which every arm has inductance, capacity and resistance.

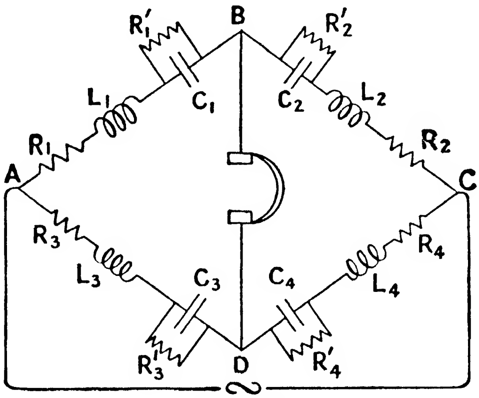


FIG. 331.

Also there is a resistance in parallel with each capacity, which may be a leak in the condenser itself. A want of balance is detected by a telephone receiver, or by some form of vibration galvanometer (p. 387).

The general equation for the condition of balance is derived from the fact that for zero current in the telephone, B and D must always be at the same potential. The equation is complex, but by giving suitable values to the electrical quantities, the equations for the most used forms of bridge may be obtained.

Let E_1 be the difference of potential between A and B, I_1 the current from A to B, and Q_1 the charge on the condenser C_1.

Then, $$E_1 = L_1\frac{dI_1}{dt} + R_1 I_1 + \frac{Q_1}{C_1}$$

and, $$I_1 = \frac{dQ_1}{dt} + \frac{Q_1}{R_1'C_1}.$$

If now, $$Q_1 = Q_0\epsilon^{jpt}$$
$$\frac{dQ_1}{dt} = jpQ_0\epsilon^{jpt} = jpQ_1$$

and, $$I_1 = jpQ_1 + \frac{Q_1}{R_1'C_1}, \text{ or, } Q_1 = \frac{I_1}{jp + \frac{1}{R_1'C_1}}$$

[1] Oliver Heaviside, *Phil. Mag.*, Feb. 1887, p. 173.

Also, $$I_1=I_0\epsilon^{j(pt+\theta)}$$

$$\frac{dI_1}{dt}=jpI_0\epsilon^{j(pt+\theta)}=jpI_1$$

$$\therefore\ E_1=I_1\left\{L_1jp+R_1+\frac{1}{C_1jp+\dfrac{1}{R_1'}}\right\}$$

with similar expressions for E_2, E_3 and E_4.

Applying Kirchhoff's first law to the points B and D, we see that for zero current in BD, $I_1=I_2$, and $I_3=I_4$.

For B and D to be at the same potential, $E_1=E_3$ and $E_2=E_4$,

$$\therefore\ \frac{E_1}{E_3}=\frac{E_2}{E_4}$$

$$\therefore\ \frac{L_1jp+R_1+\dfrac{1}{C_1jp+\dfrac{1}{R_1'}}}{L_3jp+R_3+\dfrac{1}{C_3jp+\dfrac{1}{R_3'}}}=\frac{L_2jp+R_2+\dfrac{1}{C_2jp+\dfrac{1}{R_2'}}}{L_4jp+R_4+\dfrac{1}{C_4jp+\dfrac{1}{R_4'}}}$$

As an example it is seen at once that if the inductances are all zero and the capacities infinite, this equation reduces to $\dfrac{R_1}{R_3}=\dfrac{R_2}{R_4}$, the ordinary Wheatstone relation (p. 66).

Again, if $L=0$ throughout, and $R_1=R_3=0$, $R_1'=R_3'=\infty$, and $C_2=C_4=\infty$, we have

$$\frac{\dfrac{1}{C_1jp}}{\dfrac{1}{C_3jp}}=\frac{R_2}{R_4},\quad\text{or,}\quad\frac{C_1}{C_3}=\frac{R_4}{R_2}$$

which is the equation for de Sauty's arrangement (p. 330).

Also, if all the capacities are infinite, and $L_2=L_4=0$,

$$\frac{L_1jp+R_1}{L_3jp+R_3}=\frac{R_2}{R_4}.$$

Equating the real parts of this equation, we have $R_1R_4=R_2R_3$, and from the imaginary terms,

$$L_1jpR_4=L_3jpR_2,\quad\text{or,}\quad\frac{L_1}{L_3}=\frac{R_2}{R_4},$$

which is the condition for comparison of inductances (p. 329).

Maxwell's Bridge.[1]—The earliest form of inductance bridge is due to Maxwell, the value of an inductance L_4 being compared

[1] Maxwell, "Electricity and Magnetism," vol. ii.

with that of a capacity C_1. The arrangement of the bridge is shown in Fig. 332. It will be seen that $L_1=L_2=L_3=0$, $C_2=C_3=C_4=\infty$, and $R_1=0$. The general equation then reduces to

$$\frac{\dfrac{1}{C_1jp+\dfrac{1}{R_1'}}}{R_3}=\frac{R_2}{L_4jp+R_4}$$

$$L_4jp+R_4=R_2R_3C_1jp+\frac{R_2R_3}{R_1'}.$$

Equating the real terms, we have $R_4=\dfrac{R_2R_3}{R_1'}$, which is the Wheatstone relation for balance with steady current. On equating the imaginary terms,

$$L_4=R_2R_3C_1.$$

This gives the ratio of the inductance to the capacity when the bridge is balanced for both steady and varying current. Since

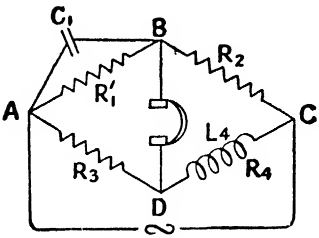


FIG. 332.

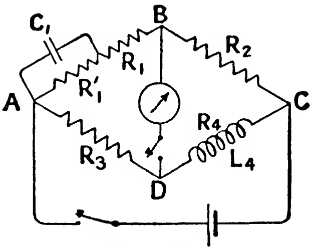


FIG. 333.

the quantity p does not appear in either equation, the balance is independent of the frequency of the applied electromotive force. In the original Maxwell method the bridge is first balanced by using a cell or battery as source of E.M.F. and closing the battery key first; then the galvanometer key. When a balance is attained, $R_2R_3=R_1'R_4$. If now, on closing the galvanometer key first and then the battery key, there is no ballistic throw, then $L_4=R_2R_3C_1$. But if the ballistic balance is not perfect, the ratio R_3/R_4 must be varied, and the double balancing repeatedly performed until both balancing conditions are attained, or an alternating E.M.F. with telephone or vibration galvanometer may be used, but for a perfect balance both conditions must be fulfilled.

Rimington's Bridge.[1]—In order to avoid the troubles in repeated double balancing of the Maxwell method, Rimington employed a form of bridge shown in Fig. 333. Instead of con-

[1] E. C. Rimington, *Phil. Mag.* (Ser. 5), **24**, p. 54. 1887.

necting the capacity in parallel with the whole resistance of the arm AB, one end is joined to A and the other end is movable. The adjustment consists in finding a position for the movable contact, such that no current flows from B to D whether steady or varying current is employed.

In this case, $L_1=L_2=L_3=0$, $C_2=C_3=C_4=\infty$.

Then from the general equation (p. 381),

$$\frac{R_1+\dfrac{1}{C_1jp+\dfrac{1}{R_1'}}}{R_3}=\frac{R_2}{L_4jp+R_4}$$

$$L_4jpR_1+R_1R_4+\frac{L_4jp+R_4}{C_1jp+\dfrac{1}{R_1'}}=R_2R_3$$

$$-L_4C_1p^2R_1+R_1R_4C_1jp+L_4jpR_1/R_1'+R_1R_4/R_1'+L_4jp+R_4$$
$$=R_2R_3C_1jp+R_2R_3/R_1'.$$

Equating real terms,

$$-L_4C_1p^2R_1+R_4(R_1+R_1')/R_1'=R_2R_3/R_1'$$

$$L_4C_1=\frac{R_4(R_1+R_1')-R_2R_3}{p^2R_1R_1'} \quad . \quad . \quad . \quad . \quad \text{(i)}$$

If steady E.M.F. is used, $p=0$, and the Wheatstone condition $R_4(R_1+R_1')=R_2R_3$ is fulfilled. But if p is not zero, then this condition cannot be fulfilled when a balance with alternating E.M.F. is attained, for we should have $L_4C_1=0$.

Equating imaginary terms,

$$R_1R_4C_1jp+L_4jpR_1/R_1'+L_4jp=R_2R_3C_1jp$$

$$\frac{L_4}{C_1}=\frac{(R_2R_3-R_1R_4)R_1'}{R_1+R_1'} \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

From equations (i) and (ii), L_4 and C_1 can be calculated, but it is necessary to use E.M.F. of one frequency only, and that frequency must be known, as p appears in equation (i).

In Rimington's original method with cell and galvanometer, $p=0$, so that from equation (i) we have the Wheatstone condition fulfilled. It is shown in the original paper that the total charge passing through the galvanometer on make or break is then zero, so that an ordinary ballistic balance is obtained when equation (ii) holds. Substituting $R_2R_3=R_4(R_1+R_1')$ in (ii), we have,

$$\frac{L_4}{C_1}=\frac{R_4(R_1')^2}{R_1+R_1'}.$$

Owen's Bridge.—In another bridge, due to Owen,[1] a com-

[1] David Owen, *Proc. Phys. Soc. Lond.*, xxvii, Dec. 1914, p. 39.

parison of an inductance L_2 with a capacity C_3 is made, the arrangement being shown in Fig. 334. In this case $L_1 = L_3 = L_4 = 0$, $C_1 = C_2 = \infty$, and $R_3' = R_4' = \infty$.

By substitution in the general equation,

$$\frac{R_1}{\frac{1}{C_3 jp}} = \frac{L_2 jp + R_2}{R_4 + \frac{1}{C_4 jp}}$$

whence
$$R_1 R_4 + \frac{R_1}{C_4 jp} = \frac{L_2 jp}{C_3 jp} + \frac{R_2}{C_3 jp}.$$

Equating the real parts of this equation,

$$R_1 R_4 = \frac{L_2}{C_3},$$

and from the imaginary parts

$$\frac{R_1}{C_4} = \frac{R_2}{C_3}.$$

These relations are independent of the frequency of the applied E.M.F., so that any form of interrupter may be used, and a telephone employed as a detector. If R_1, C_3 and C_4 are chosen, R_2 is varied until the sound in the telephone is a minimum. Then R_4 is varied until the sound is zero. If the balance is not perfect, R_2 is again adjusted for minimum sound, followed by readjustment of R_4. Then $L_2 = C_3 R_1 R_4$. The method is rapid, and is applicable to a wide range of measurement.

The mutual inductance of two coils may be found by joining the two in series and finding the resulting inductance $L_1 + L_2 + 2M$. On reversing one coil the resulting inductance is $L_1 + L_2 - 2M$. From these two values M may be calculated.

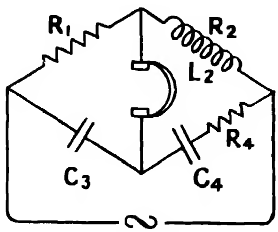


Fig. 334.

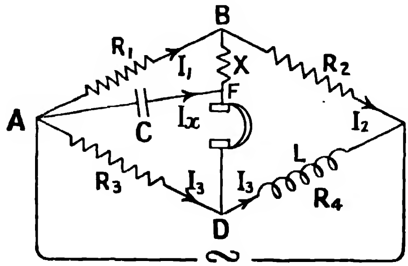


Fig. 335.

Anderson's Bridge.[1]—This is not a true Wheatstone's bridge, but as in the case of Rimington's bridge, it enables a comparison of an inductance with a capacity to be made without the repeated balancing of Maxwell's method. The scheme is shown in Fig. 335.

[1] A. Anderson, *Phil. Mag.* (Ser. 5), **31**, p. 329. 1891.

After a steady balance has been found with cell and galvanometer for the resistances R_1, R_2, R_3 and R_4, an inductive balance is obtained by varying the resistance X, placed between the junction of R_1 and R_2 and F the junction of condenser and telephone or galvanometer.

When the current in the telephone is zero, R_3 and R_4 carry the same current I_3, and $I_2=I_1+I_x$.

Then, for alternating E.M.F.'s and currents, the p.d. between A and C being E_{AC},

$$E_{AC}=(Ljp+R_3+R_4)I_3$$

and,

$$E_{AD}=R_3I_3$$

$$=\frac{R_3}{Ljp+R_3+R_4}.E_{AC}.$$

Now, for the branch ABC,

$$I_2=I_1+I_x$$

and,

$$E_{AB}=R_1I_1=\left(X+\frac{1}{Cjp}\right)I_x.$$

Also, $\quad R_1I_1+R_2I_2=E_{AC}$

$$\therefore\ E_{AC}=R_2(I_1+I_x)+R_1I_1$$

$$=(R_1+R_2)I_1+R_2I_x$$

$$=\left\{(R_1+R_2)\frac{X+\frac{1}{Cjp}}{R_1}+R_2\right\}I_x,$$

also,

$$E_{AF}=\frac{I_x}{Cjp}.$$

The condition for balance is that

$$E_{AD}=E_{AF}$$

$$\therefore\ \frac{R_3E_{AC}}{Ljp+R_3+R_4}=\frac{E_{AC}}{\left\{(R_1+R_2)\frac{\left(X+\frac{1}{Cjp}\right)}{R_1}+R_2\right\}Cjp}.$$

$$R_3(R_1+R_2)\left(X+\frac{1}{Cjp}\right)Cjp+R_1R_2R_3Cjp=R_1Ljp+R_1(R_3+R_4)$$

$$R_3(R_1+R_2)XCjp+R_3(R_1+R_2)+R_1R_2R_3Cjp$$
$$=R_1Ljp+R_1(R_3+R_4)$$

Equating real terms,

$$R_3(R_1+R_2)=R_1(R_3+R_4)$$

or,

$$R_2R_3=R_1R_4$$

which is the condition for balance with steady current.

Equating imaginary terms,

$$R_3(R_1+R_2)XC+R_1R_2R_3C=R_1L$$

$$L=C\left\{R_2R_3+\frac{R_3(R_1+R_2)}{R_1}X\right\}$$

but

$$\frac{R_1}{R_3}=\frac{R_2}{R_4}=\frac{R_1+R_2}{R_3+R_4}$$

$$\therefore L=C\{R_2R_3+(R_3+R_4)X\}.$$

If alternating E.M.F. and telephone are used, the resistances R_4 and X are alternately adjusted for minimum sound in the telephone, until a balance is obtained. In the original method the Wheatstone balance is first obtained, with cell and galvanometer. The ballistic balance for make and break is then obtained by adjustment of X.

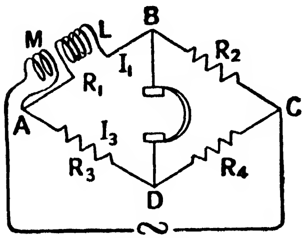


FIG. 336.

Comparison of Mutual and Self-Inductance (Maxwell [1]).—The mutual inductance of a pair of coils may be found in terms of the inductance of one of them, as shown in Fig. 336. The electromotive force produced in L on account of M, when the current grows, must be opposed to the self-inductance electromotive in L itself.

Then, when the balance has been obtained for steady current,

$$\frac{R_1}{R_2}=\frac{R_3}{R_4}.$$

Also since B and D are at the same potential,

$$R_2I_1=R_4I_3$$

or,

$$\frac{R_2}{R_4}=\frac{I_3}{I_1}$$

$$\therefore \frac{R_2+R_4}{R_4}=\frac{I_1+I_3}{I_1}.$$

The E.M.F. in AB due to self-inductance is $Ljp\mathrm{I}_1$, and that due to mutual inductance is $Mjp(\mathrm{I}_1+\mathrm{I}_3)$. Since the bridge is still balanced these must be equal and opposite.

$$\therefore LI_1=M(I_1+I_3)$$

$$\therefore LR_4=M(R_2+R_4)$$

$$\frac{L}{M}=\frac{R_2+R_4}{R_4}.$$

[1] Maxwell, "Electricity and Magnetism." vol. II.

The steady balance must be found with cell and galvanometer, but the telephone with alternating current may be used in finding the inductive balance.

Vibration Galvanometer.—In order to avoid the use of a telephone, Mr. Campbell [1] has employed a galvanometer in which a very light coil has a bifilar suspension, in which the tension can be varied by altering the pull on the suspension. The vibration frequency of the suspended coil can thus be varied (from 50 ~ to 1000 ~), and may be tuned to coincide with the frequency of the alternating current employed. Thus any feeble current in the galvanometer may, owing to resonance, produce a large vibration in the galvanometer coil, and great sensitiveness may be attained. Campbell also describes a number of methods of comparing capacities, inductances and resistance by means of the vibration galvanometer. The Einthoven string galvanometer (p. 75) may also be used for this purpose.

Comparison of Capacity and Mutual Inductance. — Campbell gives the following method for the measurement of capacity in terms of a mutual inductance. Let C be the capacity, and M (Fig. 337) be the mutual inductance, of which the coil in the circuit between D and F has inductance L, the arrangement being as shown in the figure. Since there is to be no current in the galvanometer, we have for the circuit AFD,

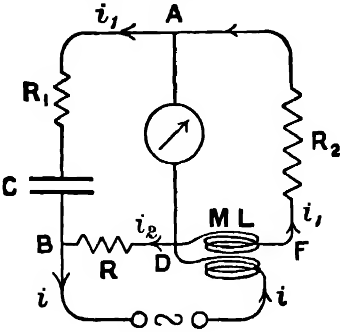


FIG. 337.

$$L\frac{dI_1}{dt}-M\frac{dI}{dt}+R_2I_1=0.$$

Now, for the point D, $I=I_1+I_2$,

$$\therefore\ L\frac{dI_1}{dt}-M\frac{dI_1}{dt}+R_2I_1=M\frac{dI_2}{dt}.$$

Again, for circuit DBA, since D and A are always at the same potential, $\frac{Q}{C}=RI_2-R_1I_1$, where $Q=\int I_1dt$, the charge upon the condenser,

$$\therefore\ \frac{\int I_1dt}{C}=RI_2-R_1I_1,$$

$$\text{or, } R_1\frac{dI_1}{dt}+\frac{I_1}{C}=R\frac{dI_2}{dt},$$

[1] Albert Campbell, *Proc. Phys. Soc. Lond.* 20, p. 626. 1907.

substituting this value of $\frac{dI_2}{dt}$ in the above,

$$L\frac{dI_1}{dt}-M\frac{dI_1}{dt}+R_2I_1=\frac{M}{R}\left(R_1\frac{dI_1}{dt}+\frac{I_1}{C}\right),$$

or, $$\left(L-M-M\frac{R_1}{R}\right)\frac{dI_1}{dt}+\left(R_2-\frac{M}{RC}\right)I_1=0,$$

and since the electromotive force is simple harmonic, the current equation may be written in terms of symbolic operators, thus,

$$\left(L-M-M\frac{R_1}{R}\right)jpI_1+\left(R_2-\frac{M}{RC}\right)I_1=0,$$

and therefore $L-M-\frac{R_1}{R}=0$, and, $R_2-\frac{M}{RC}=0$,

$$\frac{L}{M}=\frac{R+R_1}{R}, \text{ and, } \frac{M}{C}=RR_2$$

On adjusting the resistances until there is no current in the galvanometer or telephone, the relations between the inductances and capacity are known in terms of the resistances.

This method was originally due to Carey Foster [1].

Frequency Meter or Bridge.—A form of the bridge may be devised for the measurement of the frequency of the supply.

In Fig. 331 make $R_2'=R_3'=R_4'=0$, $L_2=L_3=L_4=0$, and $R_1'=\infty$.

Then the equation on p. 381 for the bridge becomes

$$\frac{L_1jp+R_1+\dfrac{1}{C_1jp}}{R_3}=\frac{R_2}{R_4}$$

$$R_4R_1+R_4\left(L_1jp+\frac{1}{C_1jp}\right)=R_2R_3$$

$$\therefore\ R_4R_1=R_2R_3$$

and $$L_1jp+\frac{1}{C_1jp}=0$$

$$1-p^2LC=0.$$

The resistance condition may be attained with steady current, once for all. The second condition gives $p=\frac{1}{\sqrt{LC}}$ and by varying L or C or both the frequency $\frac{p}{2\pi}$ may be found.

[1] G. Carey Foster, *Proc. Phys. Soc.* 8, p. 137. 1887.

Alternating current bridge measurements have become of great importance in recent years, on account of the development of telephony and wireless telegraphy and telephony. Special methods of producing currents of pure sine form have been devised by means of alternators, and even better, by means of oscillating valves (p. 587). These pure sine currents are necessary where the frequency enters into the final equation for the balanced bridge, although a timed vibration galvanometer which is not affected by currents of frequency differing from its own renders such subsidiary frequencies innocuous. With a telephone receiver as detector, the current under these conditions must have pure sine form.

So many forms of bridge have been developed for special purposes that it is impossible to give here more than the few typical forms described above. For a full account of the alternating current bridge and its uses, the student may consult "Alternating Current Bridge Methods" by B. Hague or "Alternating Current Measurements" by David Owen.

CHAPTER XII

UNITS

Dimensions.—Throughout the whole range of Physics we are concerned with the magnitudes of various quantities and their relations to each other, and it therefore becomes of importance to examine certain laws which underlie these relationships. The most fundamental relationship is that of mere number ; quantities may be added to each other provided that they are all of one kind, but not if they are of different kinds. We see therefore that all the terms that are to be added together in any equation must be of one kind, and if their nature is for any purpose changed, all the terms must change in the same manner and at the same stage of the calculation.

In order to define any physical quantity, two statements are necessary ; we must know the unit in which the quantity is measured, and the numeric relation between the quantity and the unit. The latter is a mere number or ratio, which tells us the relative magnitudes of the quantity and the unit, while the former gives us information with respect to the nature of the quantity.

We therefore require as many different kinds of unit as there are physical quantities to be measured, but the units need not necessarily be independent of each other. Before the importance of devising a scientific system of units was realised, it was customary to fix a new arbitrary unit for every fresh quantity to be measured, quite irrespectively of its relation to the units already in existence, and sometimes many units for the same quantity, as may easily be realised by contemplating the number of different units of volume there are in use in this country at the present time.

The attempt is always made in scientific work to have as few arbitrary units as possible, and to choose those units to be of as durable and easily copiable a form as possible. The fundamental units chosen are those of mass, length and time. The unit of mass is one-thousandth part of the mass of a piece of platinum kept in the Archives de Paris, and is called the gramme ; the unit of length is one-hundredth of the distance between two marks on a platinum bar at the standard temperature, also kept

at the Archives de Paris, and is called the centimetre; and the unit of time is called the second. It is $\frac{1}{86400}$ of the average interval between two successive transits of the sun across a given meridian.

Most physical units may be explicitly defined in terms of these three, raised to various powers; and the powers to which they must be raised to obtain any derived unit are called the *dimensions* of that unit. Thus the unit of volume is that of a cube whose edge is one centimetre, or writing [L] for the unit of length and [V] for the unit of volume—

$$[V]=[L^3],$$

is the dimensional equation corresponding to the above statement. It tells us that a volume is of the third dimension in length.

Or again, the unit of velocity is such that the body moves through a distance of one centimetre in one second, which fact written as a dimensional equation is—

$$[\text{Velocity}]=\left[\frac{L}{T}\right]=[LT^{-1}].$$

In a similar manner we may see that—

$$\begin{aligned}
[\text{Acceleration}]&=[LT^{-2}],\\
[\text{Force}]&=[MLT^{-2}]\\
[\text{Pressure}]&=[ML^{-1}T^{-2}],\\
[\text{Energy}]&=[F\,.\,L]=[ML^2T^{-2}],\\
[\text{Moment of inertia}]&=[ML^2],\\
[\text{Density}]&=[ML^{-3}],\\
[\text{Angle}]&=[L\,.\,L^{-1}]=[L^0].
\end{aligned}$$

An angle is of no dimensions, that is, it is a mere ratio of two lengths. These two lengths are, however, measured in different directions, and if it is desired to retain them in the dimensional equation they may be written L_x and L_y, in which case the equation becomes—

$$[\text{angle}]=[L_xL_y^{-1}]$$

In the same way—

$$[\text{couple}]=[ML_xL_yT^{-2}].$$

Knowing the units in which any quantity is to be measured, it only remains to state the numeric defining the ratio of the magnitude of the quantity to that of the unit, in order to define completely the quantity. Thus if we state that a force is $12[MLT^{-2}]$ we mean that the force is 12 units, or 12 times the force that would produce unit acceleration in unit mass. Or again, if we say that a density is $3[ML^{-3}]$ we mean that it is

3 times the unit density, that is three times the density of a substance in which there is one unit of mass in unit volume.

On the centimetre-gramme-second (C.G.S.) system, some of these derived units have particular names. Thus the unit of force is called the *dyne*, and the unit of work the *erg*.

Uses of Theory of Dimensions.—A consideration of the dimensions of the terms in a given equation frequently serves as a useful check upon the accuracy of the calculations by which the equation was obtained, since all the terms that are added in a given expression must be of the same kind, and therefore of the same dimensions.

Thus in the equation $v^2 = u^2 - 2ugt \sin \alpha + g^2t^2$, for the velocity of a body projected with velocity u at an angle α to the horizontal,

$$[v^2] = [L^2T^{-2}],\ [u^2] = [L^2T^{-2}],$$
$$[2ugt \sin \alpha] = [LT^{-1} \,.\, LT^{-2} \,.\, T] = [L^2T^{-2}],$$
and, $$[g^2T^2] = [L^2T^{-4} \,.\, T^2] = [L^2T^{-2}],$$

and we see that every term has the same dimensions. If this were not the case we should be sure of the existence of some error in the equation.

Another use to which a knowledge of dimensions may be put, is the solving of certain physical problems, thus—

Given that the difference of pressure, p, between the gas inside and outside of a soap-bubble depends only on the surface tension of the film and its radius of curvature, to find how these quantities enter into the expression for p.

The dimensional equation is $[p] = [t^x r^y]$, where x is the unknown power to which the surface tension t is to be raised, and similarly y is the unknown power of r, the radius of curvature.

Now, from p. 391, $[p] = [ML^{-1}T^{-2}]$, and t is a force per unit length, therefore $[t] = [ML^0T^{-2}] = [MT^{-2}]$, and $[r] = [L]$.

$$\therefore\ [ML^{-1}T^{-2}] = [MT^{-2}]^x[L]^y,$$
$$= [M^xL^yT^{-2x}].$$

Since these two expressions must be of the same kind—

$$x = 1, \text{ and, } y = -1,$$
$$\therefore\ [p] = [tr^{-1}],$$

that is, the pressure varies directly as the surface tension and inversely as the radius of curvature.

Again, if we are given that the velocity of a compression wave in air depends only on the pressure and density of the air,

$$[V] = [P^xD^y] = [ML^{-1}T^{-2}]^x[ML^{-3}]^y,$$
$$[LT^{-1}] = [M^{x+y}L^{-x-3y}T^{-2x}].$$

And again, since these quantities must be of the same kind—

$$x + y = 0,\ -x - 3y = 1, \text{ and, } -2x = -1,$$

from any two of which equations we have—

$$x=\tfrac{1}{2},\ \text{and},\ y=-\tfrac{1}{2},$$

$$\therefore\ [\mathrm{V}]=[\mathrm{P}^{\frac{1}{2}}\mathrm{D}^{-\frac{1}{2}}]=\left[\sqrt{\frac{\mathrm{P}}{\mathrm{D}}}\right].$$

A treatment of this kind will never give us the numerical relation between the quantities considered, as their magnitude has deliberately been excluded from the equations.

Electrical Units.—There are two equations which rest upon experiment and from these our knowledge of the dimensions of electrical quantities is derived. Taking first the expression for the force between electrical charges, namely, $\mathrm{F}=\mathrm{A}\frac{q_1q_2}{kr^2}$, the quantity A is a mere numeric and may be chosen as we please. k refers to the medium in which the charges are situated and is known as the dielectric constant. In the electrostatic C.G.S. system of units A is chosen to be unity, and on the Heaviside-Lorentz system it is chosen to be $\frac{1}{4\pi}$.[1] Only the electrostatic system will be considered here. Taking then A as unity, $\mathrm{F}=\frac{q_1q_2}{kr^2}$. In this equation F and r^2 can be expressed in centimetres, grammes and seconds, but neither q nor k can be so expressed. The only fact that is definite is that $\frac{q^2}{k}$ has the dimensions of $[\text{Force}\times r^2]=[\mathrm{ML^3T^{-2}}]$. That is, $[q]=[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{3}{2}}\mathrm{T}^{-1}k^{\frac{1}{2}}]$ or $[k]=[q^2\mathrm{M^{-1}L^{-3}T^2}]$. If k is considered to be unity for empty space then q may be called unity when F is one dyne and r one centimetre, as on p. 112, the charges being situated in a vacuum. To emphasize this, the value of k for empty space is usually written k_0.

Similar reasoning may be applied to magnetic poles, and from the equation $\mathrm{F}=\mathrm{A}\frac{m_1m_2}{\mu r^2}$, we may choose the numeric A to be unity, when the ordinary electromagnetic system of units follows. Thus $\frac{m^2}{\mu}$ has the value dynes$\times$cm.^{2} and the dimensions of m are given by $[m]=[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{3}{2}}\mathrm{T}^{-1}\mu^{\frac{1}{2}}]$ and of μ by $[\mu]=[m^2\mathrm{M^{-1}L^{-3}T^2}]$, but neither μ nor m can be expressed in centimetres, grammes and seconds only. Again, in order to emphasize the fact that poles are defined in terms of the force between them when situated in empty space, the symbol μ_0 is employed.

Relation between the Two Systems.—So long as electrostatic

[1] **For further information see "Theoretical Physics," vol. II, by W. Wilson.**

effects and magnetostatic effects are studied separately there is no relation to be obtained between q and m or between k and μ. But directly the effect of electricity in motion, that is current, upon a pole is considered, it follows that there must be some relation between these quantities. It was established first by Faraday, that an electric current is a charge in motion, and it follows that $i=\frac{dq}{dt}$, so that $[i]=[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}k^{\frac{1}{2}}][T^{-1}]$.

Now consider the interaction of a current and a pole, say the force on a pole due to a long straight current (p. 231) and the experimental fact that the value of this force is independent of the medium in which the pole and the current are situated; $F=A\frac{2im}{r}$. Remembering that $[m]=[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu^{\frac{1}{2}}]$ the force equation may be written dimensionally,

$$[MLT^{-2}]=[A][M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}k^{\frac{1}{2}}][L^{-1}][M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu^{\frac{1}{2}}]$$
$$[L^{-1}T]=[A][k^{\frac{1}{2}}\mu^{\frac{1}{2}}]$$

If then A is a constant number

$$[k^{\frac{1}{2}}\mu^{\frac{1}{2}}]=[L^{-1}T]=\left[\frac{1}{c}\right]$$

where c has the dimensions of a velocity.

On confining our attention to empty space the value of this velocity is $\frac{1}{\sqrt{k_0\mu_0}}$, a quantity which can be determined experimentally, as described on p. 398. It follows that, $k_0=\frac{1}{c^2\mu_0}$ or $\mu_0=\frac{1}{c^2k_0}$ and $A=\frac{1}{c\sqrt{k_0\mu_0}}$.

When the choice $k_0=1$ is made, this determines at once that $\mu_0=\frac{1}{c^2}$ and $A=\frac{1}{c\sqrt{\frac{1}{c^2}}}=1$, and the system of units resulting is the electrostatic system.

On the other hand, if we like to choose that $\mu_0=1$, then $k_0=\frac{1}{c^2}$ and $A=1$. This is the electromagnetic system.

A system of mixed units is sometimes employed, in which electrical quantities are given in electrostatic units and magnetic quantities are given in electromagnetic units. Equations containing electrical quantities only will be unchanged, as will those containing magnetic quantities only. But an equation containing both kinds of quantity such as $F=A\frac{2im}{r}$ will take a new

form. For the quantity i will involve k_0, and m will involve μ_0. If then we choose to make $k_0=1$ and $\mu_0=1$, it follows that $A=\frac{1}{c}$, and the equation becomes $F=\frac{1}{c}\frac{2im}{r}$. Such a system is known as the Gaussian system.

Electromagnetic System of Units.—Starting with the magnetic pole as defined above, whose dimensions are $[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu_0^{\frac{1}{2}}]$, we can derive the others from this.

Strength of Field.—Since force on magnetic pole is the product of strength of field and strength of pole (p. 3)—

$$F=Hm,$$
$$\therefore\ [MLT^{-2}]=[H][M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu_0^{\frac{1}{2}}],$$
whence,
$$[H]=[M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}].$$

This unit of magnetic field is called the *Gauss*. It is a field of such strength that a unit pole situated in it experiences a force of one dyne.

The International Congress of 1934 recommended the substitution of the name Oersted for Gauss as the name of the unit magnetic field. *Gauss* is, however, still in quite general use in this country.

Magnetic Induction.—The magnetic induction is defined on p. 234 as the quantity μH, and its dimensions are therefore $[M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu_0^{\frac{1}{2}}]$. The 1934 recommendation for the unit was *Gauss.*

Magnetic Flux (Bs).—The dimensions immediately follow from those of magnetic induction since $[s]=[L^2]$; they are,

$$[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu_0^{\frac{1}{2}}]$$

The unit on the C.G.S. system is called the *Maxwell.*

Magnetic Moment.—This may be obtained from the definition, pole strength × length, or from the couple exerted on the magnet situated in a field (p. 5). Either definition leads to the quantity—

$$[M^{\frac{1}{2}}L^{\frac{5}{2}}T^{-1}\mu_0^{\frac{1}{2}}]$$

Intensity of Magnetisation is magnetic moment per unit volume—

$$[M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu_0^{\frac{1}{2}}]$$

Electric Current.—From the relation between current and magnetic field (p. 48)

$$H=\frac{idl\,.\sin\theta}{r^2}$$

$$[i]=[M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}L]$$
$$=[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}],$$

since $\sin\theta$ is of zero dimensions. Or from the equivalence

between a current and a magnetic shell (p. 237), we have for the strength of shell $\mu_0 i$, the magnetic moment per unit area—

$$\therefore\ [\mu_0 i]=[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu_0^{\frac{1}{2}}L^{-2}]$$
$$=[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu_0^{\frac{1}{2}}]$$
$$[i]=[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}]$$

Quantity of Electricity.—Since $i=\frac{q}{t}$, or $q=it$ (p. 120)

$$[q]=[M^{\frac{1}{2}}L^{\frac{1}{2}}\mu_0^{-\frac{1}{2}}]$$

Electromotive Force.—(e) The rate of working in units of work per second is equal to the product of current and electromotive force (p. 55).

$$[i\,.\,e]=[ML^2T^{-3}]$$
$$[M^{\frac{1}{2}}L^{\frac{1}{2}}\,.\,T^{-1}\mu_0^{-\frac{1}{2}}\,.\,e]=[ML^2T^{-3}]$$
$$\therefore\ [e]=[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}\mu_0^{\frac{1}{2}}]$$

Electric Intensity.—From the relation $e=\int E dl$ (p. 119), we have—

$$[E]=[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-2}\mu_0^{\frac{1}{2}}].$$

Resistance.—The ohmic relation between electromotive force and current (p. 56) gives

$$[r]=\frac{[e]}{[i]}=[LT^{-1}\mu_0].$$

Neglecting the dimensions of μ_0, resistance is seen to be of the dimensions of a velocity, and for this reason it is sometimes spoken of as so many centimetres per second.

Capacity.—From the equation $c=\frac{q}{e}$ (p. 147) we obtain—

$$[c]=[L^{-1}T^2\mu_0^{-1}].$$

Inductance.—Using the relation $e=-l\frac{di}{dt}$, or $e=-m\frac{di}{dt}$ (p. 318), we see the dimensions of inductance, either self or mutual, to be—

$$\frac{[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}\mu_0^{\frac{1}{2}}][T]}{[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}]}=[L\mu_0]$$

Again neglecting the dimensions of μ_0, an inductance may be measured in centimetres.

Electrostatic System of Units.—Beginning with the unit of electrical charge, $[q]=[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}k_0^{\frac{1}{2}}]$, as defined on p. 393, we may obtain the other electrical and magnetic units in terms of this.

Potential Difference.—As defined on p. 120, we have—

$$\text{p.d.}\times\text{charge}=\text{work},$$
$$\therefore\ [e\,.\,q]=[ML^2T^{-2}],$$

from which,

$$[e]=\frac{[ML^2T^{-2}]}{[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}k_0^{\frac{1}{2}}]},$$
$$=[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}k_0^{-\frac{1}{2}}]$$

Electric Intensity.—Since force on a charge is equal to product of charge and electric intensity (p. 112)—

$$[\mathrm{E}q]=[\mathrm{MLT}^{-2}],$$
$$\therefore\ [\mathrm{E}]=[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{-\frac{1}{2}}\mathrm{T}^{-1}k_0^{-\frac{1}{2}}].$$

Electrical Induction, Displacement, or Surface Density.—From the definition on p. 128, this is equal to $k\mathrm{E}$.

$$\therefore\ [\sigma]=[\phi]=[\mathrm{D}]=[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{-\frac{1}{2}}\mathrm{T}^{-1}k_0^{\frac{1}{2}}].$$

Electric Current.—The rate at which electric charge passes along a conductor is the current.

$$\therefore\ [i]=\frac{[q]}{[\mathrm{T}]}=[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{3}{2}}\mathrm{T}^{-2}k_0^{\frac{1}{2}}].$$

Resistance.—As on p. 396—

$$[r]=\frac{[e]}{[i]};$$
$$\therefore\ [r]=\frac{[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{1}{2}}\mathrm{T}^{-1}k_0^{-\frac{1}{2}}]}{[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{3}{2}}\mathrm{T}^{-2}k_0^{\frac{1}{2}}]}$$
$$=[\mathrm{L}^{-1}\mathrm{T}k_0^{-1}].$$

Magnetic Field.—Using again the expression $\mathrm{H}=\dfrac{idl\sin\theta}{r^2}$, we have—

$$[\mathrm{H}]=[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{1}{2}}\mathrm{T}^{-2}k_0^{\frac{1}{2}}].$$

Magnetic Pole.—Since, force$=m\mathrm{H}$—

$$[m]=[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{1}{2}}k_0^{-\frac{1}{2}}].$$

Capacity.—Since, $c=\dfrac{q}{e}$—

$$[c]=\frac{[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{3}{2}}\mathrm{T}^{-1}k_0^{\frac{1}{2}}]}{[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{1}{2}}\mathrm{T}^{-1}k_0^{-\frac{1}{2}}]}=[\mathrm{L}k_0].$$

Inductance.—From the definition $e=-l\dfrac{di}{dt}$, or $e=-m\dfrac{di}{dt}$—

$[m]$[1] $=[l]=\dfrac{[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{1}{2}}\mathrm{T}^{-1}k_0^{-\frac{1}{2}}][\mathrm{T}]}{[\mathrm{M}^{\frac{1}{2}}\mathrm{L}^{\frac{3}{2}}\mathrm{T}^{-2}k_0^{\frac{1}{2}}]}=[\mathrm{L}^{-1}\mathrm{T}^2k_0^{-1}].$

Relation between Units on the Two Systems.—It was seen on p. 394 that the quantity $\dfrac{1}{\sqrt{k_0\mu_0}}$ is of the nature of a velocity and it is desirable to find the value of this velocity. This may be done by comparing the magnitudes of any one electrical quantity on the two systems, the electrostatic and the electromagnetic. The value found is very nearly 3×10^{10} cm. per second, which is

[1] *m* is here mutual inductance, not magnetic pole.

the velocity of light. This suggestive fact led Maxwell to conclude that light consists of an electromagnetic wave, and he eventually established the fact by means of equations relating to the electrical and magnetic condition of the field.

If i_e be the number of electrostatic units in a given current, the complete expression for the current is $i_e[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}k_0^{\frac{1}{2}}]$, and if i_m be the number of electromagnetic units in the same current, $i_m[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}]$ is its expression in electromagnetic measure, where i_e and i_m are mere numbers.

$$\therefore\ i_e[M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}k_0^{\frac{1}{2}}]=i_m[M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}],$$

or,
$$\left[\frac{1}{k_0^{\frac{1}{2}}\mu_0^{\frac{1}{2}}}\right]=\frac{i_e}{i_m}[LT^{-1}].$$

But i_e and i_m being the magnitudes of the same current in different units, their ratio is the inverse ratio of the size of the units,

$$\therefore\ \frac{i_e}{i_m}=\frac{\text{size of electromagnetic unit of current}}{\text{size of electrostatic unit of current}}=\text{say } c.$$

We see, then, that $\dfrac{1}{\sqrt{k_0\mu_0}}=c$ centimetres per second, since $[LT^{-1}]$ is a velocity of one centimetre per second.

The numerical value of c may be determined by measuring experimentally the same current in electrostatic and in electromagnetic measure. It is, however, more convenient to choose capacity for the subject of measurement, as the capacity of a condenser of simple form may be calculated in electrostatic measure from its dimensions, and it may be measured in electromagnetic measure by means of the ballistic galvanometer.

Let a given condenser have a capacity of c_e electrostatic units, or c_m electromagnetic units.

Then, as before, $\quad c_e[Lk_0]=c_m[L^{-1}T^2\mu_0^{-1}],$

or,
$$\left[\frac{1}{k_0\mu_0}\right]=\frac{c_e}{c_m}[L^2T^{-2}],$$

$$\left[\frac{1}{\sqrt{k_0\mu_0}}\right]=\sqrt{\frac{c_e}{c_m}}[LT^{-1}],$$

$$\therefore\ \frac{1}{\sqrt{k_0\mu_0}}=\sqrt{\frac{c_e}{c_m}}=c \text{ centimetres per second.}$$

A convenient form of condenser may be made by fixing layers of tinfoil upon two sheets of glass, one of the layers being circular, and surrounded by a circular guard ring, the other covering the whole sheet; or from two pieces of silvered plate glass from which the paint on the back of the silver has been removed by

means of caustic soda, and a circular gap made in one of them by scraping away the silver. The two are placed with the metallic surfaces face to face and kept apart by three thin distance-pieces of ebonite, the thickness of which will give the distance apart of the plates of the condenser (t). Then $c_e=\frac{A}{4\pi t}$, where A is the area of the circular plate, the larger plate being earthed. The dielectric constant of air is taken as unity (Fig. 338).

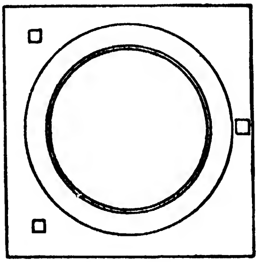


FIG. 338.

To determine c_m, the plate and ring are charged to a high potential V; then the ring is earthed, and the plate discharged through the ballistic galvanometer.

$$c_m V=\frac{cT}{2\pi AH}\theta, \qquad \text{(p. 257)}$$

where θ is the ballistic throw. The galvanometer may be calibrated by producing a steady deflection θ_1 by means of a current produced by a known fraction of V, say $\frac{V}{n}$ and a high resistance r.

$$\frac{VAH}{nR}=c\theta_1,$$

$$\therefore\ c_m=\frac{T}{2\pi nR}\cdot\frac{\theta}{\theta_1},$$

If the capacity is so small that an unreasonably high potential V is required to produce a readable ballistic throw, the capacity may be compared with that of a larger condenser by the method on p. 162, or in terms of a resistance and a frequency, by the method on p. 400.

The principle of the first method was employed by Professors Ayrton and Perry,[1] the condenser being charged by the fall of potential over a resistance of 10,000 ohms produced by a battery of 382 Daniell cells. To produce the steady current in the galvanometer, a known fraction of this was used, and a high resistance was placed in series with the galvanometer. The mean of their results, corrected for the value of the B.A. ohm used by them, in terms of the international ohm is, $c=2{\cdot}995\times10^{10}$.

Maxwell's Method.—If a condenser be placed in series with a battery and galvanometer, it will receive a charge ec, where e is the electromotive force of the battery, and c the capacity of the condenser. This is the state of affairs when the rocker D is in

[1] W. E. Ayrton and J. Perry, *Journal Soc. Tel. Eng.*, 8, p. 126. 1879.

contact with A (Fig. 339). Then if D is moved over so that it makes contact with B instead of A, the condenser is discharged. On moving D back into contact with A the condenser receives another charge ec, and if this process be repeated n times per second, the total charge that has been drawn from the battery and which has passed through the galvanometer is nec. This is equivalent to a current, and if n is great in comparison with the frequency of vibration of the moving part of the galvanometer, a steady deflection will be obtained. The key may take the form of a revolving commutator or a suitably arranged tuning-fork of known frequency, in which case n is known.

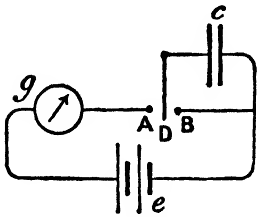


FIG. 339.

If the condenser and key are replaced by a conductor, and the whole resistance of the circuit adjusted until the deflection of galvanometer is the same as that with the condenser and key, the current

$$i=\frac{e}{r}=nec.$$

$$\therefore\ r=\frac{1}{nc},\ \text{or},\ c=\frac{1}{rn}.$$

We see, therefore, that the intermittent charge and discharge has the same effect as a resistance, and if the frequency and the whole resistance of the circuit are known, the capacity c may be determined.

Since the capacity has been found in electromagnetic measure, and its value in electrostatic measure can be calculated from its dimensions, the velocity c can be found as before.

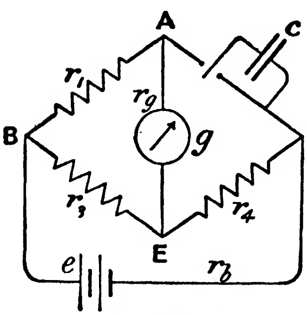


FIG. 340.

Maxwell pointed out [1] that the substitution of the resistance for the capacity and key is unnecessary if these are placed in one arm of the Wheatstone's bridge and a balance obtained in the ordinary way.

The arrangement is then as shown in Fig. 340. The resistances are adjusted until the galvanometer deflection is zero, when the approximate relation $\frac{1}{nc}\cdot r_3=r_1r_4$ holds. Since the method is

[1] Maxwell, "Electricity and Magnetism," vol. ii, §§ 775 and 776.

usually employed for measuring very small capacities, r_3 is always very small. For example, a condenser of the type described on p. 399 would have a capacity of the order 10^{-22} absolute electromagnetic unit or 10^{-13} farad.

If, then, n is say 100, and r_1 and r_4 say 1,000,000 ohms each—

$$r_3 \cdot \frac{1}{10^{-11}} = 10^{12}$$

$$r_3 = 10 \text{ ohms.}$$

In a case such as this the relation between capacity and resistance may be established, by equating the values of the steady current in the galvanometer when the condenser circuit is permanently open, the value of which is $\dfrac{e}{r_b + r_4 + \dfrac{r_3(r_1+r_g)}{r_3+r_1+r_g}} \cdot \dfrac{r_3}{r_3+r_1+r_g}$, where e is the electromotive force of the battery, to the charge per second passing through the galvanometer due to the intermittent charging of the condenser. The difference of potential between A and D is,

$$(r_4 \times \text{current in ED}) + (r_g \times \text{current in } g)$$

$$= r_4 \frac{e}{\mathrm{P}} + r_g \frac{e}{\mathrm{P}} \cdot \frac{r_3}{r_3+r_1+g} = \frac{e}{\mathrm{P}} \cdot \left(r_4 + \frac{r_3 r_g}{r_3+r_1+r_g}\right),$$

where P is written for $\left\{r_b + r_4 + \dfrac{r_3(r_1+r_g)}{r_3+r_1+r_g}\right\}$, the resistance of the entire circuit exclusively of the branch AD. Hence charge on C when fully charged is—

$$\frac{ce}{\mathrm{P}}\left(r_4 + \frac{r_3 r_g}{r_3+r_1+r_g}\right).$$

Now, owing to the smallness of the resistances r_3 and r_b, the charge, when the condenser is closed, will flow round the circuit A(BE)eD, B and E being practically one point, owing to the smallness of r_3. The charge divides between the paths AB and AEB, the fraction $\dfrac{r_1}{r_3+r_1+r_g}$ flowing by the path AEB, that is, through the galvanometer, the ratio being independent of the inductances of the branches. And since this discharge takes place n times per second, the current in the galvanometer due to this cause is—

$$\frac{nce}{\mathrm{P}} \cdot \left(r_4 + \frac{r_3 r_g}{r_3+r_1+r_g}\right) \cdot \frac{r_1}{r_3+r_1+r_g};$$

therefore when the galvanometer deflection is zero—

$$\frac{e}{\mathrm{P}}\cdot\frac{r_3}{r_2+r_1+r_g}=\frac{nce}{\mathrm{P}}\left(r_4+\frac{r_3r_g}{r_3+r_1+r_g}\right)\cdot\frac{r_1}{r_3+r_1+r_g},$$

$$r_3=ncr_1\left(r_4+\frac{r_3r_g}{r_3+r_1+r_g}\right).$$

But the last fraction is negligible since r_3 is very small in comparison with r_1, and therefore $nc=\frac{r_3}{r_1r_4}$.

For a complete discussion when no restrictions are placed on the magnitudes of the resistances, the student is referred to "Absolute Measurements in Electricity and Magnetism," by A. Gray. The expression there obtained for nc is—

$$\frac{r_3\{(r_3+r_4+r_b)(r_g+r_1+r_3)-r_3^2\}}{\{r_1(r_3+r_4+r_b)+r_3r_b\}\ \{r_4(r_g+r_1+r_3)+r_gr_3\}},$$

which reduces to the above when r_1 and r_4 are very great in comparison with the other resistances.

Employing this method and using a spherical condenser, E. B. Rosa [1] found c to be $3{\cdot}0004\times10^{10}$.

Sir J. J. Thomson and Dr. G. F. C. Searle [2] used a cylindrical condenser provided with guard rings of cylindrical form at the ends, which necessitated a slight modification of the bridge connections. The mean of their values for c is $2{\cdot}9955\times10^{10}$.

The value of c has also been found by measuring a capacity in terms of an inductance and two resistances (see p. 382), and also in terms of a resistance and a time by means of the slow discharge of a condenser (see p. 317). Another interesting method is to determine the frequency of oscillatory discharge when a condenser discharges through a known resistance and inductance (see p. 338), the frequency being found by obtaining a photograph of the spark upon a revolving photographic plate. In this way Lodge and Glazebrook [3] found $c=3{\cdot}009\times10^{10}$.

The latter determinations give values differing very slightly from each other. There is little doubt that the value of $\frac{1}{\sqrt{k_0\mu_0}}$ is very nearly $3{\cdot}00\times10^{10}$ cm. per second, which is also the velocity of light in empty space.

Practical Units.—We have described in various places (see pp. 56 and 313) the manner in which the practical units are chosen in order that they may be of convenient sizes, while retaining simple relationships with the absolute electromagnetic

[1] E. B. Rosa, *Phil. Mag.* (Ser. 5), **28**, p. 315. 1889.
[2] J. J. Thomson and G. F. C. Searle, *Phil. Trans.*, **181**, p. 583. 1890.
[3] O. J. Lodge and R. T. Glazebrook, *Cambr. Phil. Trans.*, **18**, p. 136. 1899.

units. Thus the ampere is one-tenth of the absolute C.G.S. electromagnetic unit of current, and the volt is 10^8 absolute units of electromotive force. From these are derived the ohm, the joule and the watt, which are respectively 10^9, 10^7, and 10^7 times the corresponding absolute units. Similarly the farad is the capacity of a conductor which is raised in potential by one volt by a charge of one coulomb, and hence it is equal in value to $\frac{10^{-8}}{10}=10^{-9}$ absolute units. This unit is still very large for practical purposes, so a millionth of it, called a micro-farad is usually employed; its value is therefore 10^{-15} absolute unit. Again, the henry is the inductance of a circuit in which a rate of change of current of one ampere per second is accompanied by an electromotive force of 1 volt, and it is therefore $\frac{10^8}{10^{-1}}=10^9$ absolute units. Owing to its inconvenient size, the milli-henry, or thousandth of a henry, is usually employed in practice. Its value is 10^6 absolute units.

The above electrical and magnetic units are collected into the following table—

Unit.	Dimensions.		Ratio of electro-magnetic to electro-static unit.	Practical unit.	
	Electrostatic.	Electromagnetic.		Name.	Ratio of size to that of electro-magnetic unit.
Quantity	$M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}k_0^{\frac{1}{2}}$	$M^{\frac{1}{2}}L^{\frac{1}{2}}\mu_0^{-\frac{1}{2}}$	c	Coulomb	10^{-1}
Electromotive force	$M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}k_0^{-\frac{1}{2}}$	$M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}\mu_0^{\frac{1}{2}}$	c^{-1}	Volt	10^8
Electric intensity .	$M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}k_0^{-\frac{1}{2}}$	$M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-2}\mu_0^{\frac{1}{2}}$	c^{-1}		
Electric displacement	$M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}k_0^{\frac{1}{2}}$	$M^{\frac{1}{2}}L^{-\frac{3}{2}}\mu_0^{-\frac{1}{2}}$	c		
Electric current . .	$M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}k_0^{\frac{1}{2}}$	$M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}$	c	Ampere	10^{-1}
Electric resistance .	$L^{-1}Tk_0^{-1}$	$LT^{-1}\mu_0$	c^{-2}	Ohm	10^9
Capacity	Lk_0	$L^{-1}T^2\mu_0^{-1}$	c^2	Farad	10^{-9}
Inductance . . .	$L^{-1}T^2k_0^{-1}$	$L\mu_0$	c^{-2}	Henry	10^9
Magnetic field . .	$M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-2}k_0^{\frac{1}{2}}$	$M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu_0^{-\frac{1}{2}}$	c	Gauss or Oersted	1
Magnetic induction .	$M^{\frac{1}{2}}L^{-\frac{3}{2}}k_0^{-\frac{1}{2}}$	$M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu_0^{\frac{1}{2}}$	c^{-1}		
Magnetic flux . .	$M^{\frac{1}{2}}L^{\frac{1}{2}}k_0^{-\frac{1}{2}}$	$M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu_0^{\frac{1}{2}}$	c^{-1}	Maxwell	1
Magnetic moment .	$M^{\frac{1}{2}}L^{\frac{3}{2}}k_0^{-\frac{1}{2}}$	$M^{\frac{1}{2}}L^{\frac{5}{2}}T^{-1}\mu_0^{\frac{1}{2}}$	c^{-1}		
Intensity of magnetisation	$M^{\frac{1}{2}}L^{-\frac{3}{2}}k_0^{-\frac{1}{2}}$	$M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu_0^{\frac{1}{2}}$	c^{-1}		
Magnetic pole . .	$M^{\frac{1}{2}}L^{\frac{1}{2}}k_0^{-\frac{1}{2}}$	$M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu_0^{\frac{1}{2}}$	c^{-1}		
Dielectric constant .	k_0	$L^{-2}T^2\mu_0^{-1}$	c^{-2}		
Magnetic permeability	$L^{-2}T^2k_0^{-1}$	μ_0	c^2		
Energy	ML^2T^{-2}	ML^2T^{-2}		Joule	10^7
Rate of working .	ML^2T^{-3}	ML^2T^{-3}		Watt	10^7

International Units.—The need for reproducible standards for everyday measurements led in 1861 to the appointment of a Committee by the British Association for the Advancement of Science and various specifications were issued starting with one for the ohm, which was defined in terms of the resistance of a mercury column of stipulated length, cross-section and temperature. The "B.A. ohm" was in turn largely replaced in 1884 by a closer approximation, the "legal ohm" (authorised in Great Britain by Order in Council); the "international ohm" followed, adopted in 1908 at the International Conference on Electrical Units and Standards.

Absolute determination of current was not so difficult, using a tangent galvanometer or a current balance, but the voltameter was made the basis of the international ampere, which liberates 0·00111800 gm. per sec. of silver in a suitable voltameter. The international volt follows from these, but the mercury standard of resistance was awkward to realise in everyday work, so the international volt was stated in terms of the electromotive force of the Weston cell, this E.M.F. being at first given as 1·0184 volts (1908), later (1910) revised to 1·0183 volts.

The divergence between the international units and the quantities they were meant to represent are now known with some precision and from 1948, by international agreement, the international units are discarded in favour of the true units. With an uncertainty of about 1 in the last figure, the ratio of the international to the true unit is as follows: ohm, 1·00049; ampere, 0·99985; volt, 1·00034. The "international joule," defined as the rate of working with 1 international ampere flowing in a resistance of 1 international ohm, is 1·00019 absolute joule.

M.K.S. Systems of Units.—Several systems of units have been proposed in which the fundamental units are the metre, the kilogramme, the second and one electrical quantity. Unit acceleration in such systems is 1 m./sec.^{2}, and unit force, which gives a mass of 1 kgm. this acceleration, is the *newton*, which is $10^2 \times 10^3 = 10^5$ dynes. This leads to the unit of work as the newton-metre or $10^5 \times 10^2 = 10^7$ ergs, the *joule*. If one of the practical units, say the ohm, is then taken as a primary standard, the other practical units follow. When some concrete approximation, such as the international ohm, is adopted, not one of the practical electrical units will be a "true" unit as originally intended in its definition.

In the M.K.S. System (proposed by Giorgi in 1904 and modified since) the magnetic permeability of free space $\mu_0 = 10^{-7}$ approximately. By making this relation exact, the "true" units can be retained. In this system the dielectric constant of free space $k_0 = 10^{11}$.

The systems so far considered have laid emphasis on the action between point charges, regarded as the basis of the subject.

There is a powerful movement, among electrical engineers especially, to transfer the emphasis to fields and to "rationalise" the units, so defining them that the 4π which appears in a number of formulæ disappears, only of course to reappear in others in which the advocates are less interested. In the system adopted in 1950 at the Paris meeting of the International Electrotechnical Commission, $\mu_0=4\pi/10^7$ and $k_0=(1/36\pi)\times10^{-9}$. In this system the capacity of a parallel-plate condenser is kA/d (A in sq. m., d in metres), but the "permittivity" k here is the ordinary dielectric constant divided by $(36\pi\times10^9)$. The dimensions and the derived units are also different. For example, magnetic field strength is measured in amperes (or ampere-turns) per metre. A great number of papers have appeared in the last few years on the relative merits etc. of the rival systems.[1]

Determination of Practical Standards.—The following pages give some account of the most important steps which established the practical standards referred to above. The current unit has usually been determined by some form of current balance in series with a voltameter and the result expressed in terms of silver deposited. Most of the methods used for the resistance unit depend on a comparison of the potential difference across the conductor passing a known current with the E.M.F. developed in a conductor rotating in a magnetic field.

The "legal ohm" (1894) (p. 404) was defined as the resistance to steady current of a column of mercury of uniform cross-section, having a length of 106·300 cm. and a mass of 14·4521 gms. at 0° C. In 1908 the London Electrical Conference revised this by stipulating a cross-sectional area of 1 sq. mm. and stating that there should be spherical end-pieces, full of mercury, of 4 cm. diameter provided with sealed-in platinum current and potential leads arranged in a specified manner and a formula was provided representing the additional resistance to be added to allow for these ends.

Determination of the Ohm.—(i) *Rotating Coil.* The Committee of the British Association in 1863 adopted, for constructing a standard of resistance, the method of rotating a closed coil of wire about a vertical axis in the earth's magnetic field, the deflection of a magnetic needle suspended at the centre of the coil being observed.

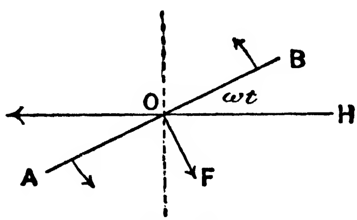


FIG. 341.

If AB in Fig. 341 is the plan of the circular coil when its plane makes angle ωt with the magnetic meridian, $\pi a^2 n\mathrm{H}\sin\omega t$ is the magnetic flux passing through the coil, where a is its radius, n the number of turns, H the horizontal component of the earth's

[1] See, for example, G. F. Nicholson, *Brit. Journ. App. Phys.*, 2, p. 177 (1951); Sir Charles Darwin, *Nature*, 164, p. 262 (1949).

magnetic field, and ω the angular velocity of rotation about the vertical axis O. The momentary electromotive force round the coil is—

$$-\pi a^2 n \mathrm{H}\frac{d}{dt}(\sin \omega t) = -\pi a^2 n \mathrm{H}\omega \cos \omega t,$$

and since this is an alternating electromotive force of maximum value, $-\pi a^2 n\mathrm{H}\omega$, we know from p. 344 that the momentary current i is—

$$-\frac{\pi a^2 n \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} \cos (\omega t-\theta),$$

since the current is angle θ in phase behind the electromotive force, where $\tan \theta = \frac{l\omega}{r}$, l being the inductance of the coil and r its resistance.

This current gives rise to a magnetic field OF whose value at the centre is $\frac{2\pi n i}{a}$ (p. 49). The component of this at right angles to the meridian is—

$$\frac{2\pi n i}{a} \cos \omega t = -\frac{2\pi^2 a n^2 \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} \cos (\omega t-\theta) \cos \omega t,$$

$$= -\frac{2\pi^2 a n^2 \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} (\cos^2 \omega t \cos \theta + \tfrac{1}{2} \sin 2\omega t \sin \theta).$$

The mean value of $\cos^2 \omega t$ for a complete cycle we have seen (p. 347) to be $\frac{1}{2}$, and the mean value of $\sin 2\omega t$ is zero, therefore mean magnetic field at right angles to the meridian is—

$$-\frac{\pi^2 a n^2 \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} \cos \theta.$$

In a similar manner we see that the instantaneous component of the field in the meridian is,

$$\frac{2\pi n i}{a} \sin \omega t = -\frac{2\pi^2 a n^2 \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} \cos (\omega t-\theta) \sin \omega t$$

$$= -\frac{2\pi^2 a n^2 \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} (\tfrac{1}{2} \sin 2\omega t \cos \theta + \sin^2 \omega t \sin \theta),$$

the mean value of which is—

$$-\frac{\pi^2 a n^2 \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} \sin \theta.$$

The resultant field in the meridian is therefore,

$$\mathrm{H} - \frac{\pi^2 a n^2 \mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}} \sin \theta,$$

and the suspended needle will then be in equilibrium when making an angle ϕ with the meridian such that—

$$\frac{\dfrac{\pi^2an^2\mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}}\cos\theta}{\mathrm{H}-\dfrac{\pi^2an^2\mathrm{H}\omega}{\sqrt{l^2\omega^2+r^2}}\sin\theta}=\tan\phi.$$

Since H occurs in every term on the left-hand side of this equation it disappears, and we see that the equilibrium position of the needle is independent of its value;

$$\frac{\pi^2an^2\omega\cos\theta}{\sqrt{l^2\omega^2+r^2}-\pi^2an^2\omega\sin\theta}=\tan\phi;$$

$\tan\phi$ therefore depends upon the velocity of revolution ω. This must of course be so great that the separate impulses acting on the needle follow at intervals sufficiently small in comparison with the period of vibration of the needle for the deflection to be steady. If the inductance of the coil is small enough for the quantity $l\omega$ to be negligible in comparison with r, $\sin\theta=0$ and $\cos\theta=1$, and we then have—

$$r=\pi^2an^2\omega\cot\phi.$$

The angular velocity ω having the dimensions of the inverse of a time, and a being a length, we see that r has the dimensions of a velocity, and its determination depends upon the accurate measurement of these two quantities, together with an angle ϕ.

It will be noticed that the effect of the torsion in the suspension fibre, and the influence of the magnetic field of the suspended needle in inducing current in the rotating coil, have been omitted. These must be measured and allowed for. Standard resistances constructed by comparison with the coil whose resistance was determined in absolute measure by this means were distributed by the British Association.

Lord Rayleigh,[1] in 1882, made a determination of the ohm by this method. The inductance of the coil was calculated from its dimensions and also determined by the method on p. 327. The velocity of rotation of the coil was determined by the stroboscopic method. He found that,

1 B.A. unit $=0{\cdot}98651$ earth quadrant per second.

A rotating coil method due to W. Weber,[2] in which the coil is turned through 180° in the earth's field, the current passing through a ballistic galvanometer and the throw being noted, has also been used by him and by G. Wiedemann,[3] the latter of whom

[1] Lord Rayleigh, *Phil. Trans.*, **173**, p. 661. 1882.
[2] W. Weber, *Pogg. Ann.*, **82**, p. 337. 1851.
[3] G. Wiedemann, *Abhandl. Berlin Akad. d. Wiss.*, 1884.

found the ohm to be the resistance of a column of mercury 106·162 cm. long, 1 sq. mm. in cross-section at 0° C. The method is similar in principle to that of the earth inductor described on p. 264.

Determination of the Ohm.—(ii) *Method of Lorenz.*[1]—The movement of a conductor in a magnetic field gives rise to an electromotive force which is equal to the rate at which magnetic flux is being cut by the conductor. If, then, a conducting disc of radius a is rotated with constant angular velocity, n times per second, when its plane is at right angles to a magnetic field of strength H, any radius of the disc cuts a flux $\pi a^2 \mathrm{H}$ in each revolution, and therefore the electromotive force acting from the axis to the circumference is $\pi a^2 n \mathrm{H}$. If the field is produced by a current i in a pair of circular coils co-axial with the disc (shown by a dotted circle in the diagram), $\pi a^2 \mathrm{H}$ becomes mi, where m is the mutual inductance of the coils and the disc, and therefore, electromotive force is equal to nmi. This electromotive force is balanced against the difference of potential between the ends of a resistance r (Fig. 342) in series with the coils and through which the current i is flowing. When the galvanometer G is undisturbed—

$$ri = nmi,$$

or,

$$r = nm.$$

Lord Rayleigh and Mrs. Sidgwick[2] carried out a measurement of the ohm by this method in 1883, but instead of employing a

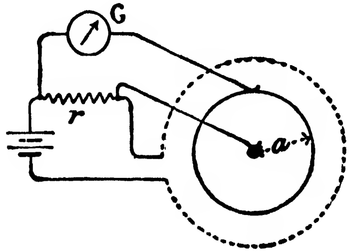


Fig. 342.

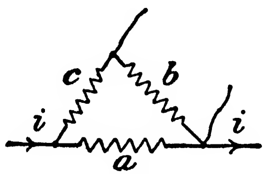


Fig. 343.

calibrated tube of mercury for the resistance r, they used three wire resistances, a, b and c (Fig. 343), of which a is the smallest and carries most of the current i, while c is large compared with b. The fall of potential over b is balanced against the electromotive force in the rotating disc. If then i be the total current in the fixed coils (Fig. 342), that in b and c is $\dfrac{ai}{a+b+c}$, and the

[1] L. Lorenz, *Pogg. Ann.*, **149**, p. 251. 1873.
[2] Lord Rayleigh and Mrs. Sidgwick, *Phil. Trans.*, **174**, p. 295. 1883.

difference of potential between the ends of b is $\frac{abi}{a+b+c}$. The resistance a consisted of two unit coils in parallel, b was a platinum-silver $\frac{1}{10}$ unit, and c was, in three series of experiments, 10, 16 and 20 respectively. The value of b in terms of the absolute unit is then given by the experiment, and the mercury column that has the unit resistance can therefore be found if the specific resistance of mercury in terms of the B.A. unit, in which b is known can be found. This forms the subject of another paper by Lord Rayleigh and Mrs. Sidgwick.[1]

To return to the Lorenz method, the rotating disc is of brass, and has a diameter of 31·072 cm.; and contact was made with its edge by a copper brush well amalgamated, the contact at the shaft being of a similar kind and touching it on a circle whose diameter is 2·096 cm. The mean distance apart of the mean planes of the two coils is 3·275 cm. for two series of experiments, and 30·6944 cm. for a third. From careful measurements of the coils, the value of m for the first two series of experiments is found to be 214·569, and for the third 110·392. In the first series of experiments the speed of the disc was about 12·8 revolutions per second, in the second 8, and in the third 12·8, and was determined by the stroboscopic method, the standard being a calibrated tuning-fork.

The general scheme of connections is shown in Fig. 344. In order to eliminate the electromotive forces in the galvanometer circuit due to thermo-electric effects at the sliding contacts, and the cutting of the earth's vertical magnetic field by the disc as it rotates, a small difference of potential is maintained between the points A and B, connected by a low resistance which is adjusted until the galvanometer reading remains constant with the disc running, but without the main current, whether the galvanometer circuit is broken or closed. In the actual experiment r is not adjusted to give an exact balance, but some value such as r_1 is employed, and the difference of the galvanometer readings when the main current is reversed by means of the key K is observed. r_1 is then changed to the value r_2, such that the galvanometer deflection for either position of K is the reverse to that in the previous

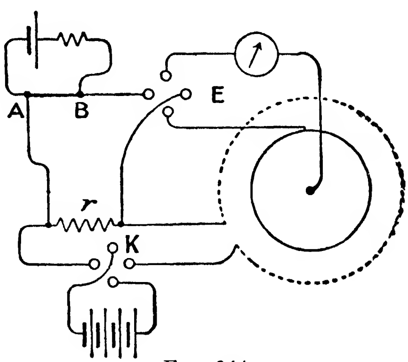


FIG. 344.

[1] Lord Rayleigh and Mrs. Sidgwick, *Phil. Trans.*, 174, p. 193. 1883.

case, and the difference in deflection for a reversal again noted. By interpolation the value of r for an exact balance is calculated.

The mean of the results gave the value of the B.A. unit to be $0{\cdot}98677 \times 10^9$ C.G.S. units.

Determination of the Ohm.—(iii) *Mutual Inductance of Two Coils.*—Kirchhoff [1] suggested a method, the principle of which has been given on p. 327. Using a galvanometer of the suspended magnet type, the throw when current i is started in the primary coil is given by—

$$\frac{mi}{r_1}=\frac{HT}{\pi G}\sin\tfrac{1}{2}\theta,$$

and when a steady current is maintained in the secondary by the fall of potential across a small resistance r, in the primary—

$$\frac{irG}{r_1 H}=\tan\theta_1$$

$$\therefore\ m=\frac{rT}{\pi}\cdot\frac{\sin\frac{1}{2}\theta}{\tan\theta_1},$$

or,

$$r=m\frac{\pi\tan\theta_1}{T\sin\frac{1}{2}\theta}.$$

If then m is determined by calculation, in absolute measure, and T, θ and θ_1 observed, r is known.

This method has been used by Rowland, Glazebrook, and others. As a mean of his results, Glazebrook [2] has given that—

$$1\text{ B.A. unit}=0{\cdot}98665\times10^9\text{ C.G.S. units.}$$

The method of finding the heat produced in a wire by means of the calorimeter was employed by Joule; and also the method of damping (see p. 261) due to W. Weber has been employed for determining the magnitude of the ohm in terms of the resistance of a mercury column, but the results are not so consistent, nor are the methods capable of such accuracy as the above.

Determination of the Electro-chemical Equivalent of Silver.—With the object of defining the C.G.S. unit of current in terms of some quantity that may be conveniently reproduced, Lord Rayleigh and Mrs. Sidgwick [3] made a determination of the electro-chemical equivalent of silver. The form of voltameter employed has already been described (p. 64). The current is passed for a measured time, about three-quarters of an hour, through two or three such voltameters in series, and through a system of coaxial coils shown in Fig. 345. The smaller coil is

1 G. Kirchhoff, *Pogg. Ann.*, 76, p. 412. 1849.
2 Glazebrook, *B. A. Report*, p. 97. 1890.
3 Lord Rayleigh and Mrs. Sidgwick, *Phil. Trans.*, 175, p. 411. 1884.

suspended from the beam of a balance, and the force upon it due to the currents is found by taking the difference in the weighings when the current in the larger coils, which are fixed, is reversed. We know that the potential energy of a coil carrying current i_2 due to another carrying current i_1 is mi_1i_2, where m is the mutual inductance of the coils (p. 319). Hence the force on the small coil in the direction of the axis is $i_1i_2\frac{dm}{dx}$, where x is in the direction of the axis. The actual calculation of the force between the coils is beyond the scope of this book, and the student who wishes for more information is advised to consult the original paper. It may be noted that the small coil is placed in such a position that the force on it is a maximum, and hence does not vary appreciably for a small change in its position, thus obviating the necessity of determining the relative positions of the coils with any high degree of accuracy.

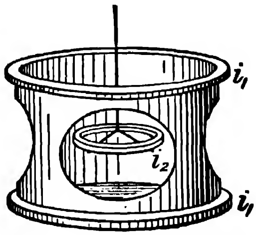


FIG. 345.

The best results are obtained when a solution of pure silver nitrate in water is used in the voltameter, and with a 3-inch platinum bowl and a solution of strength 15 to 30 per cent., a current of 1 ampere may be passed for an hour.

As a mean result it was found that the C.G.S. unit of current deposits 0·0111794 gramme of silver per second.

Sir F. E. Smith and Prof. T. Mather [1] found, in 1908, that the ampere deposits 0·00111827 gramme of silver per second.

Standards of Electromotive Force.—In performing the above-mentioned work on the electro-chemical equivalent of silver, Lord Rayleigh and Mrs. Sidgwick at the same time found the electromotive force of the Clark cell (p. 189).

The method is essentially that of the potentiometer. A battery of two Leclanché cells, B (Fig. 346), maintains steady current in the two resistance boxes C and D, and the electromotive force of the Clark cell e balanced against the fall of potential over the box D. The total resistance of the two boxes C and D is maintained constant so that the current i in them is constant. Then $e=r_1i$, where r_1 is the resistance in D when the galvanometer G indicates zero current. The potential difference between the ends of the resistance r which carries the current of the electrolysis experiment (above), and passes through the voltameter V, is now included in the circuit of e, and if P is this potential difference, the resistance r_2 of D which is now required to balance the electromotive force $e-$P, is r_2i.

[1] F. E. Smith and T. Mather, *Phil. Trans.*, Ser. A, **207**, p. 546. **1908.**

$$\therefore \frac{e-P}{e}=\frac{r_2}{r_1},$$

and,
$$P=e\left(1-\frac{r_2}{r_1}\right),$$

or if the current i in r is known from the weighing experiment—

$$ri=e\left(1-\frac{r_2}{r_1}\right).$$

The mean result indicated that the electromotive force of the Clark cell at 15° C. is 1·435 volt.

The Clark cell was for a time the basis of the legal definition of the volt, for an Order in Council dated 24 August, 1894, defined the volt as 0·6974 (*i.e.* 1/1·434) of the terminal potential difference of a Clark cell at 15° C. The Clark cell was displaced by the Weston cell as the standard (see p. 189) because of the short life of a Clark cell, the zinc alloying with one of the platinum

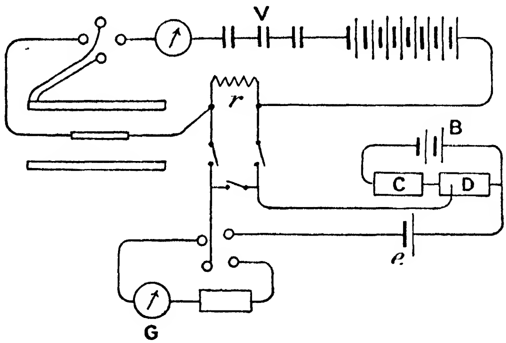


FIG. 346.

leads, causing it to swell and crack the glass. Weston cells, constructed properly from pure materials and used carefully, have very long lives and are very stable, and also have a small temperature-coefficient. The E.M.F.s of cells made to the same specification by different workers may disagree by some 20 or 30 microvolts.

CHAPTER XIII

ELECTROMAGNETIC RADIATION

Fundamental Equations.—The state of a field of electric and magnetic force at any point may be represented by means of three fundamental relations, already dealt with in their general forms, which it is now our purpose to express in reference to rectangular co-ordinates.

(i) *Gauss's Law.*—The total normal electrical flux over a closed surface is equal to 4π times the total charge within it (p. 123). A similar relation holds in the magnetic case (p. 234). Let us consider the electric intensity E at a point A (Fig. 347) to be resolved into components parallel to the three axes of co-ordinates, the components being P parallel to OX, Q and R parallel to OY and OZ respectively. Then if P varies from point to point as we move from A in a direction parallel to OX, at the rate $\frac{dP}{dx}$, its value at the face BD of the very small rectangular solid ABCD will be $P+\frac{dP}{dx}dx$.

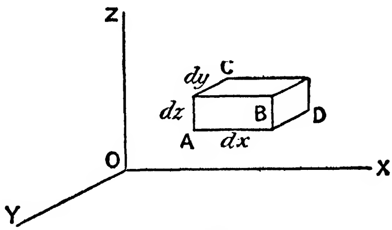


FIG. 347.

The normal flux over the face AC is $kP \,.\, dy\, dz$, where k is the dielectric constant and $dy \,.\, dz$ the area of the face AC. The normal flux over BD is $k\left(P+\frac{dP}{dx}dx\right)dy\ dz$, and the difference between these two is the contribution of the two faces AC and BD to the total normal flux over the whole surface. That is—

$$k\left(P+\frac{dP}{dx}dx\right)dy\ dz-kPdy\ dz=k\frac{dP}{dx}dx\ dy\ dz.$$

Treating the faces AB and CD in the same way, we get $k\frac{dQ}{dy}dx\ dy\ dz$ as the contribution to the normal flux for these faces, and similarly $k\frac{dR}{dz}dx\ dy\ dz$ that for the faces AD and BC.

If now there is a volume density of electric charge ρ, the amount of charge within the surface ABCD is $\rho dx\, dy\, dz$, and thus from Gauss's law we have—

$$k\left(\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}\right)dx\, dy\, dz=4\pi\rho dx\, dy\, dz,$$

or,

$$\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}=\frac{4\pi\rho}{k}.$$

For the magnetic field, if α, β and γ are the components of H, the equation is—

$$\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}=\frac{4\pi\rho}{\mu}$$

where ρ is the volume density of magnetic pole, and μ the magnetic permeability of the medium. This is Poisson's equation.

If $\rho=0$, we have for the electric field—

and for the magnetic field,

$$\left.\begin{aligned}\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}&=0\\ \frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}&=0\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad \text{(i)}$$

The quantity $\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}$ is frequently called the *divergence* of the vector quantity E, and is written div. E, since it represents the rate at which E increases or diminishes as we pass outwards from the point; then—

$$\text{div. } E=\frac{4\pi\rho}{k}.$$

(ii) *Line Integral of Magnetic Field round a Current.*—On p. 230 we saw that the work done in carrying a unit pole round a closed path through which a current is flowing, is equal to 4π times the current; in other words, the line integral of the magnetic field round the closed path is 4π times the current.

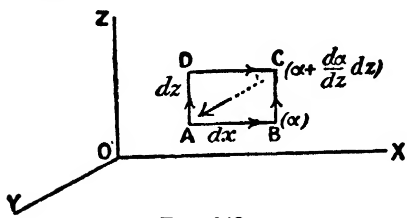


FIG. 348.

If u, v and w are the components of the current density, or current per unit area, the current flowing through the small rectangle ABCD (Fig. 348) whose plane is perpendicular to OY is $vdxdz$, the area being $dxdz$. If the value of the magnetic field along AB be α and along DC be $\alpha+\frac{d\alpha}{dz}dz$, the work done on the pole as it moves

along AB is αdx, and along CD is $-\left(\alpha+\frac{d\alpha}{dz}dz\right)dx$. Similarly, for DA it is $-\gamma dz$, and for BC $+\left(\gamma+\frac{d\gamma}{dx}dx\right)dz$.

Therefore for the whole circuital path ABCD—

$$\text{Work done } =\alpha dx-\left(\alpha+\frac{d\alpha}{dz}dz\right)dx-\gamma dz+\left(\gamma+\frac{d\gamma}{dx}dx\right)dz$$

$$=\left(\frac{d\gamma}{dx}-\frac{d\alpha}{dz}\right)dx\,dz;$$

and the above law therefore gives us—

$$4\pi v\,dx\,dz=\left(\frac{d\gamma}{dx}-\frac{d\alpha}{dz}\right)dx\,dz,$$

or,

$$-4\pi v=\frac{d\alpha}{dz}-\frac{d\gamma}{dx}.$$

Treating the other components of the current u and w in a similar manner, we get two other similar equations. The law may then be expressed by the three equations—

$$\left.\begin{aligned}-4\pi u&=\frac{d\gamma}{dy}-\frac{d\beta}{dz}\\ -4\pi v&=\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\\ -4\pi w&=\frac{d\beta}{dx}-\frac{d\alpha}{dy}\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

(iii) *Electromotive Force round Circuit through which the Magnetic Flux is varying.*—The law $e=-\frac{dN}{dt}$ (p. 251) may be expressed by means of its components in an exactly similar manner. Referring to Fig. 348, if μ is the magnetic permeability of the medium, the flux N through the rectangle ABCD is $\mu\beta dx\,.\,dz$, and $dx\,.\,dz\frac{d}{dt}(\mu\beta)$ is the rate of change of flux. If the component of electric field E along AB is P, and along DC, $\left(P+\frac{dP}{dz}dz\right)$, and the components along AD and BC respectively R and $\left(R+\frac{dR}{dx}dx\right)$, the whole electromotive force e round the rectangle is—

$$Pdx-\left(P+\frac{dP}{dz}dz\right)dx-Rdz+\left(R+\frac{dR}{dx}dx\right)dz=\left(\frac{dR}{dx}-\frac{dP}{dz}\right)dx\,dz;$$

and since $e=-\frac{dN}{dt}$, we have, remembering Rule II, p. 252—

$$\mu\frac{d\beta}{dt}=\frac{dP}{dz}-\frac{dR}{dx}.$$

With the two corresponding equations for the components α and γ, we have altogether—

$$\left.\begin{aligned}\mu\frac{d\alpha}{dt}&=\frac{dR}{dy}-\frac{dQ}{dz}\\ \mu\frac{d\beta}{dt}&=\frac{dP}{dz}-\frac{dR}{dx}\\ \mu\frac{d\gamma}{dt}&=\frac{dQ}{dx}-\frac{dP}{dy}\end{aligned}\right\} \quad \ldots\ldots\ldots \quad \text{(iii)}$$

Maxwell's Displacement Current.—The above sets of equations are the expression of certain experimentally established laws, and do not depend upon any assumption as to the nature of the mode of action occurring in the dielectric. In equations (ii) the current u, v, w, means that an electric charge is moving in a certain direction, and we know that this motion cannot be continuous unless the medium is an electrical conductor. The possibility of a current in a dielectric was pointed out by Maxwell, who, following Faraday, was bent upon explaining electromagnetic phenomena in terms of actions occurring in the dielectric. According to Maxwell, any change in the electrical induction in a medium is an electric current. Thus, when a current flows into a condenser (Fig. 349) by means of conducting leads AB and CD, the current in these leads is the rate at which charge passes on to the plates of the condenser. If, for simplicity, we assume the plates of the condenser each to have unit area, the charge on each plate is the surface density σ, and therefore the current i in the leads is given by—

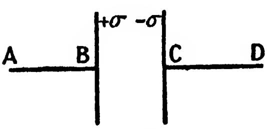


FIG. 349.

$$i=\frac{d\sigma}{dt}.$$

But the electric displacement D in the medium between the plates, we saw on p. 128 to be equal to σ;

$$\therefore\ i=\frac{dD}{dt}.$$

Hence we may consider the current to be continuous through the condenser, its value in the dielectric itself being the rate of

change of the electrical displacement; the essential difference between the dielectric and a conductor being that D cannot exceed a certain amount for each value of the electric intensity, so that after the current has flowed for a short time, D becomes constant and the current ceases.

The question may be better understood by considering an analogy. Let the condenser be replaced by a vessel having an indiarubber membrane stretched across it so as to obstruct the flow of water brought in by the pipe AB (Fig. 349) and emerging by the pipe CD. For a given difference of hydrostatic pressure between B and C, the flow continues until the displacement of the membrane reaches a certain amount, determined by its elasticity, and then ceases; but while the displacement is increasing there is an actual current of water, in through AB and out through CD.

According to Maxwell, the magnetic effect of a displacement current is similar to that of a conduction current, the distinction between the two being quite artificial. Equations (ii) therefore apply to the displacement current, which is the only current, in a dielectric.

Rowland [1] has proved that a moving "statical" charge produces magnetic effects similar to those of a conduction current, by rotating an ebonite disc having alternate sectors which were gilt and charged, and observing the deflection of a magnet, placed under and near to the disc. The direction of the deflection of the magnet was that which would be produced by an electric current corresponding to the moving charge, and was reversed on reversal of the sign of the charge, or the direction of rotation.

We may now, by means of the relation on p. 128, write the components u, v and w, of the displacement current in terms of those of electric intensity, P, Q and R, and the dielectric constant of the medium.

The general equation $D=\frac{kE}{4\pi}$, by differentiation with respect to t gives—

$$i=\frac{dD}{dt}=\frac{k}{4\pi}\cdot\frac{dE}{dt},$$

which, written in terms of its components, becomes—

$$u=\frac{k}{4\pi}\cdot\frac{dP}{dt},\quad v=\frac{k}{4\pi}\cdot\frac{dQ}{dt},\ \text{and,}\ w=\frac{k}{4\pi}\cdot\frac{dR}{dt},$$

[1] H. A. Rowland and C. T. Hutchinson, *Phil. Mag.* (Ser. 5), **27**, p. 445, 1889.

and substituting these values for u, v and w in equations (ii), we have—

$$\left.\begin{aligned}-k\frac{dP}{dt}&=\frac{d\gamma}{dy}-\frac{d\beta}{dz}\\ -k\frac{dQ}{dt}&=\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\\ -k\frac{dR}{dt}&=\frac{d\beta}{dx}-\frac{d\alpha}{dy}\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv)}$$

Equations (iii) and (iv) contain the six variable quantities α, β, γ and P, Q, R, and by eliminating five of these we can obtain an equation involving one only. Before dealing with this general equation we will treat one or two simpler cases.

Propagation of Plane Wave.—The simplest case of wave motion to consider is that of a plane wave, that is, a wave in which the electric intensity is at any instant the same over the whole plane. Let us take the plane YOZ (Fig. 350) as the plane of the wave, by which we mean that all over this plane the electric and the magnetic intensities are of constant value at any given moment. It follows that these quantities have each the same value for all values of y and z, and their variations in the Y and Z directions are zero; hence their differential coefficients with respect to y and z are likewise zero.

Equations (iii) then reduce to—

$$\mu\frac{d\alpha}{dt}=0, \quad \mu\frac{d\beta}{dt}=-\frac{dR}{dx}, \text{ and, } \mu\frac{d\gamma}{dt}=\frac{dQ}{dx},$$

and consequently α is zero or a constant. But constant values do not enter into the wave propagation, so that for our purposes we may put $\alpha=0$.

Similarly equations (iv) reduce to—

$$k\frac{dP}{dt}=0, \quad k\frac{dQ}{dt}=\frac{d\gamma}{dx}, \text{ and, } k\frac{dR}{dt}=-\frac{d\beta}{dx}.$$

Therefore $P=0$, and since we have seen that $\alpha=0$, it follows that the directions of the electric and magnetic intensities are entirely in the plane of the wave.

We may choose any direction in the plane YOZ that we please for that of the electric intensity, and we will therefore take it parallel to OZ.

Then $Q=0$, and we see from either of the relations

$$\mu\frac{d\gamma}{dt}=\frac{dQ}{dx}, \text{ or, } k\frac{dQ}{dt}=\frac{d\gamma}{dx}$$

that in this case $\gamma=0$.

Hence the electric and magnetic intensities are at right angles to each other; if R is the only component of the electric intensity, then β is the only component of the magnetic field.

The equations now reduce to

$$\mu\frac{d\beta}{dt}=-\frac{dR}{dx}, \text{ and, } k\frac{dR}{dt}=-\frac{d\beta}{dx}.$$

Differentiating the first with respect to t and the second with respect to x,

$$\mu\frac{d^2\beta}{dt^2}=-\frac{d^2R}{dxdt}, \text{ and, } k\frac{d^2R}{dxdt}=-\frac{d^2\beta}{dx^2},$$

$$\therefore \frac{d^2\beta}{dt^2}=\frac{1}{k\mu}\cdot\frac{d^2\beta}{dx^2}.$$

Or, differentiating the first with respect to x and the second with respect to t,

$$\mu\frac{d^2\beta}{dxdt}=-\frac{d^2R}{dx^2}, \text{ and, } k\frac{d^2R}{dt^2}=-\frac{d^2\beta}{dxdt},$$

$$\therefore \frac{d^2R}{dt^2}=\frac{1}{k\mu}\cdot\frac{d^2R}{dx^2}.$$

In each of the above equations the quantities may be considered to be all expressed in the same units. Or if it is desired to look upon P, Q and R as being expressed in electrostatic units the conversion into electromagnetic units may be made by writing P/c, Q/c, R/c, and $1/c^2\mu$ in place of P, Q, R and k, where c is the ratio of the electromagnetic unit to the electrostatic unit (pp. 394 and 403). On then eliminating as above, the equation $\frac{d^2\beta}{dt^2}=c^2\frac{d^2\beta}{dt^2}$ is obtained, in which c^2 has replaced $\frac{1}{k\mu}$.

The equation for R is in the form of the general equation for a plane wave, the direction of propagation being parallel to the axis OX, and its general solution is $R=f_1(x-vt)+f_2(x+vt)$, where $v^2=\frac{1}{k\mu}$ and f_1 and f_2 are any functions, and v is the velocity of the wave in the medium. It becomes c for empty space where $c^2=\frac{1}{k_0\mu_0}$. The truth of the solution may be established by differentiating this value of R twice with respect to t and twice with respect to x, and substituting the values in the differential equation. $f_1(x-vt)$ and $f_2(x+vt)$ are general expressions for wave motion along the axis of x, the former in the direction OX and the latter in the reverse direction XO. For, after a given interval of time t_1 the expression $f_1(x-vt)$ becomes $f_1(x-vt-vt_1)$, and if the origin be moved

forward along OX by the distance vt_1 the abscissæ referred to the new origin being x_1, then $x=x_1+vt_1$. The expression for the disturbance is then $f_1(x_1+vt_1-vt-vt_1)=f_1(x_1-vt)$; that is, referred to the new origin it has the same form as it had when referred to the old origin at a time t_1 earlier. Its form is therefore unchanged, but it has moved forward with velocity v.

In a similar manner the wave $f_2(x+vt)$ may be shown to travel unchanged in form and with velocity v in the opposite direction to the wave $f_1(x-vt)$.

We shall only consider the wave which travels forward, and only the most important case of such a wave, namely that in which R and β vary harmonically.

Let $R=R_0 \sin \frac{2\pi}{\lambda}(x-vt)$; then at the instant from which time is reckoned, $t=0$, and

$$R=R_0 \sin \frac{2\pi}{\lambda}x.$$

This equation gives the value of R at all points in space at this instant, and the ordinates of the curve R (Fig. 350) represent the

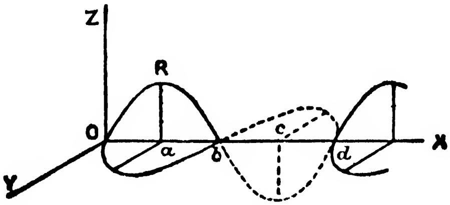


FIG. 350.

distribution of the electric intensity. Although the curve is drawn with the axis OX as axis of reference, it must be understood that the value of R at all points in any plane parallel to YOZ is the same at each instant, and is represented by the ordinate of the curve. If x be increased by the length λ,

$$R=R_0 \sin \frac{2\pi}{\lambda}(x+\lambda)=R_0 \sin \left(\frac{2\pi}{\lambda}x+2\pi\right)$$

and the curve begins to repeat itself. λ is called the wave-length, which in the diagram is the length Od. At the point b, $x=\frac{\lambda}{2}$ and $R=0$; while at a and c, x is $\frac{\lambda}{4}$ and $\frac{3\lambda}{4}$, and $R=R_0$ and $-R_0$ respectively.

Again, if the wave travels distance λ in time T, $\frac{\lambda}{T}$ is the velocity

v, and substituting this value for v in the equation for R, we have—

$$R=R_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right).$$

In order to find the expression for β, we make use of the equation

$$\mu\frac{d\beta}{dt}=-\frac{dR}{dx}.$$

$$\frac{dR}{dx}=+\frac{2\pi}{\lambda}R_0 \cos 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right),$$

$$\therefore \frac{d\beta}{dt}=-\frac{2\pi}{\mu\lambda}R_0 \cos 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right).$$

Integrating which, we get,

$$\beta=\frac{T}{\mu\lambda}R_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right).$$

But,

$$\frac{T}{\lambda}=\frac{1}{v}=\sqrt{k\mu},$$

$$\therefore \beta=\sqrt{\frac{k}{\mu}}R_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right).$$

The maximum value of β is therefore $\sqrt{\frac{k}{\mu}}R_0$, and calling this β_0, we have—

$$\beta=\beta_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right).$$

This curve also is plotted in Fig. 350.

Magnetic Field and Motion of Faraday Tubes.—The followers of Faraday and Maxwell have insisted upon the possibility of explaining electrical phenomena in terms of processes occurring in the dielectric, and in particular Sir J. J. Thomson has shown that while on the one hand electrostatic phenomena may be described in terms of tubes of induction and the stresses occurring in them, on the other hand a magnetic field is simply an attribute of the Faraday tubes in motion. In his work on "Recent Researches in Electricity and Magnetism" he has shown that the ordinary electromagnetic laws are consistent with the assumption that the motion of the ends of the Faraday tubes constitutes the electric current in the conductor along which they are moving, while in the dielectric the magnetic field is a vector quantity drawn at right angles to the tubes and to their direction of motion, whose magnitude is $4\pi Dv \sin\theta$, where θ is the angle between the tube and its direction of motion, and D is the

number of unit tubes per unit area at right angles to their direction, or as we saw on p. 128, the electric displacement,[1] and v is the velocity of the tubes. When the motion of the tubes is at right angles to their direction, $\sin \theta = 1$, and

$$H = 4\pi D v, \text{ or } H = kEv,$$

since,

$$E = \frac{4\pi D}{k}$$

If AC and BD (Fig. 351) are the two plates of a condenser of which AC is negatively charged and BD positively, then on connecting them by the wire CD, the Faraday tubes whose ends are near C and D approach each other along the conducting path now provided, and leave a space into which the neighbouring ones are pushed by their lateral pressures (p. 134). These in their turn collapse along CD, so that there is a general movement of the tubes, a few of which are indicated in the diagram, in the direction OV. This means a conduction current along BD where the positive ends of the tubes are travelling, and in AC a positive current from C to A in the opposite direction to the travel of the negative ends of the tubes. The displacement current in the dielectric is from the upper plate to the lower, since D is diminishing.

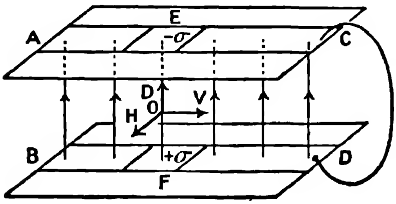


FIG. 351.

When the plates are so large that the electric field between them may be considered to be uniform, consider two strips of unit width in the direction of the currents in AC and BD. The charges upon unit areas F and E of these strips are $+\sigma$ and $-\sigma$ respectively, and these are equal to D, the electric displacement at the point O between them. If at any instant these charges are moving with velocity v, the current i in each strip is $\sigma v = Dv$. But if width of plate is b, the total current in either is ib, and the line integral of the magnetic field linked with the current is $4\pi ib$ (p. 230), and $Hb = 4\pi ib$, and therefore the magnetic field H at the point O between the plates at the given instant is $4\pi i$—

$$\therefore H = 4\pi D v.$$

It should be noticed that the direction of the displacement current is in the line of D, that is at right angles to the motion of the tubes.

[1] Sir J. J. Thomson uses the letter N to represent this quantity, but we have used D in order to emphasise the identity with Maxwell's electric displacement and to avoid confusion with N, the magnetic flux.

Plane Wave considered as Motion of Faraday Tubes.—Returning to the case of the plane wave, we may, in order to get a vivid picture of the processes occurring, replace the electrical intensity R by the displacement $\frac{kR}{4\pi}=D$, which represents the number of Faraday tubes per unit area, and the equation for the wave becomes—

$$D=D_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)$$

where, $$D_0=\frac{kR_0}{4\pi}.$$

The distribution of Faraday tubes at the time $t=0$ will therefore be somewhat as shown in Fig. 352, in which the vertical lines

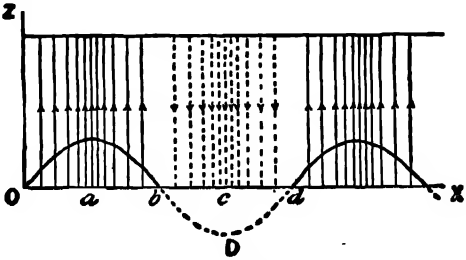


FIG. 352.

represent the tubes, their number per unit area being D, the electrical displacement. The full lines represent positive values of D and the dotted lines negative values, and they are drawn so that the closeness with which they are packed, or their number per unit length measured along OX, is proportional to the ordinate of the curve,

$$D=D_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)$$

remembering that the diagram is a section of the whole field, the direction of the axis OY being at right angles to the plane of the diagram. Every section of the field parallel to the plane ZOX would be similar to the diagram at the given instant, since the value of D over any plane parallel to YOZ is constant.

If now the whole system of tubes is imagined to be moving in the direction OX with velocity v, we shall have a representation of the harmonic plane wave travelling forwards, in accordance with Maxwell's equations.

The magnetic intensity β is everywhere at right angles to D and to v, and is therefore perpendicular to the plane of the diagram, its magnitude being given by $\beta=4\pi Dv$ (p. 422). Since

v is constant, β is at each point proportional to D, and is therefore a maximum at points a and c and zero at the points b and d. In fact, its value is given by the curve of β in Fig. 350, which fact is consistent with the result obtained on p. 418, directly from the equations of the field.

The wave that we are considering is polarised, in the optical sense, the electrical displacement being everywhere parallel to OZ, and the magnetic intensity parallel to OY. Considerations of the problem of reflection have shown that the magnetic intensity takes place in what in optics is called the plane of polarisation, the electric intensity being at right angles to it. YOX is therefore the plane of polarisation. It is not easy to see how the condition indicated in Figs. 350 and 352 can be brought about at a given place in the dielectric, but if an electric oscillation in a given direction occur at a point at a very great distance, the waves, of whatever form they may be near the point, are practically plane at great distances from it, and although the problem of the origin of the waves may present many difficulties, these are very much reduced when the waves have spread out far enough to become plane.

Energy of Wave.—The energy associated with the electric displacement D in the dielectric we have seen to be $\frac{2\pi D^2}{k}$ per unit volume (p. 130), or $\frac{kR^2}{8\pi}$, where R is the electric intensity. In the case of a plane wave, R varies harmonically at every given point of the medium, and therefore the mean value of R^2 for a cycle of change is $\frac{1}{2}R_0^2$ (p. 347), where R_0 is the maximum value of R. Hence the mean value of the energy per unit volume of the dielectric, as the wave passes through it, is $\frac{kR_0^2}{16\pi}$. Similarly, on account of the magnetic field β, the mean energy per unit volume of the medium is $\frac{\mu\beta_0^2}{16\pi}$, and the mean energy per unit volume due to both these effects is $\frac{kR_0^2+\mu\beta_0^2}{16\pi}$, and is the proper measure of the intensity of the radiation at any point.

We have seen above, that $\beta_0=\sqrt{\frac{k}{\mu}}R_0$ (p. 421), and therefore $\mu\beta_0^2=kR_0^2$. Consequently—

$$\text{energy per unit volume}=\frac{kR_0^2}{8\pi}=\frac{\mu\beta_0^2}{8\pi}=\frac{R_0\beta_0}{8\pi c}.$$

The energy of the wave is therefore half of it associated with

the electrical intensity or displacement, the other half being associated with the accompanying magnetic field.

Poynting's Theorem.—An extremely important theorem on the transfer of energy, due to Prof. Poynting,[1] throws a great deal of light upon the propagation of electromagnetic waves, and also upon the flow of energy when a current is passing in a conductor. On p. 422 we saw that in a plane wave, the magnetic intensity β is related to the electric displacement D, by the equation $\beta=4\pi Dv$, or since $D=\frac{kR}{4\pi}$, $\beta=kRv$, where β and R are at right angles to each other and to the direction of propagation, v being the velocity whose value is $\frac{1}{\sqrt{k\mu}}$. From Fig. 351 on p. 422 it will be seen that directions of β, R and v are related in the manner given by the left-hand law on p. 240. Now, the energy per unit volume is $\frac{\mu\beta^2}{8\pi}$, due to the magnetic field, and $\frac{kR^2}{8\pi}$ due to the electrical field, and the sum of these two is—

$$\frac{\mu\beta^2+kR^2}{8\pi}=\frac{\mu\beta kRv+kR\,.\,\frac{\beta}{kv}}{8\pi}=\frac{\beta R}{4\pi v},$$

from above.

Since the condition is travelling forwards with velocity v, energy is streaming past the point considered at the rate $\frac{\beta R}{4\pi v}\,.\,v=\frac{\beta R}{4\pi}$ units per second through each unit of area of cross-section of the wave at any instant. In all other cases of the transference of energy, as, for example, the flow of heat, the energy entering a given space may be measured by the amount passing through the boundary of the space. Prof. Poynting suggested that the above may be the expression of a very general law, the direction of the flow of energy being determined by the directions of the electrical and magnetic intensities, its value being proportional to their product. If the direction of one of the quantities β or R be reversed, the direction of the flow of energy, that is, of the wave propagation, is reversed; a conclusion which we shall arrive at independently on p. 440.

The theorem may be applied to several important cases:—

(i) The steady current i in a wire is accompanied by a dissipation of energy at the rate ie units per unit length of wire, where e is the fall of potential per unit length of the wire. In the air immediately surrounding the wire e is also the electric intensity.

[1] J. H. Poynting, *Phil. Trans. Roy. Soc.*, **175**, p. 343. 1884.

Further, the magnetic intensity H is given by $\frac{2i}{r}$, where r is the radius of the wire. Now e is directed along the wire, and H at right angles to it, and the quantity $\frac{eH}{4\pi}$ therefore represents a flow of energy from the dielectric into the wire, paying regard to the left-hand law mentioned on the last page. Thus the rate at which energy enters unit length of the wire is $\frac{eH}{4\pi}\times 2\pi r$ ergs per second, where $2\pi r$ is the area of surface of unit length of wire. And since $H=\frac{2i}{r}$,

$$\text{flow of energy into wire} = \frac{e}{4\pi}\times 2\pi r \,.\, \frac{2i}{r} = ei \text{ ergs per second,}$$

which is the rate of dissipation of energy in the form of heat in the wire.

It is therefore reasonable to suppose that the energy from the source of supply does not travel along the wire, but through the dielectric, entering the wire through its lateral surface. It should be noted that if the direction of e be reversed, that of i and of H are both reversed, and the direction of propagation of energy is still from the dielectric into the wire.

(ii) The slow discharge of a condenser affords another example of the application of the law. Let the wire ABC join the plates of the condenser (Fig. 353), and for simplicity let the direction of the wire be everywhere parallel to the electrical intensity. As the current flows in the wire from A to B energy enters the wire as in case (i), to be there dissipated as heat. But the electric displacement in the dielectric itself is directed from the plate EA towards FC, and the magnetic field is from front to back (see Fig. 351), which would indicate, according to the law, a flow of energy parallel to the plates of the condenser from EF to AC, and at a point such as P it is travelling in the direction of the arrow towards the wire.

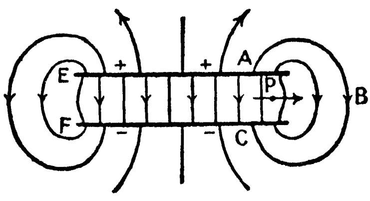


FIG. 353.

(iii) The current from a battery flows along the wire joining the poles; but we have seen that this consists of the positive ends of the Faraday tubes travelling from the positive pole, and the negative ends from the negative pole, the two approaching each other along the conductor as the tubes contract. The function of the battery is to furnish a continuous supply of Faraday tubes,

and their mode of disappearance is similar to that in the case of the discharge of the condenser. The energy flows through the dielectric, the wire merely acting as a means of directing the motion of the ends of the tubes and converting their energy into heat.

(iv) In the case of an electromagnetic wave whose direction of propagation is parallel to a conducting surface, as in Fig. 354, the dissipation of energy in the form of heat in the conductor is zero if the conductivity is infinite, the flow of energy being in the direction of propagation of the wave, none of it entering the conductor. But if the conductor has resistance, the motion of the ends of the Faraday tubes along it involves the expenditure of energy. That is, energy enters the conductor, and hence the electric intensity and the Faraday tubes must be inclined to the surface of the conductor where they meet it. Application of the left-hand law now indicates that the direction of inclination is as shown in the figure, and it should be noticed that the magnetic intensity is parallel to the conducting surface, but from back to front at A, where the electric intensity R is positive, and from front to back at B, where R is negative. The resistance of the conductor has the effect of retarding the ends of the Faraday tubes, which are therefore dragged along by the rest of the tube in opposition to this retardation.

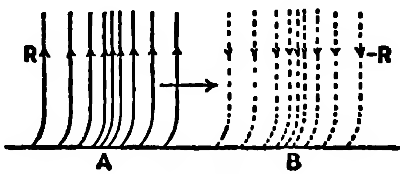


FIG. 354.

(v) An alternating electromotive force causes energy to enter the conductor whatever the direction of the electromotive force, as we saw in example (i); but during the first half-period the electromotive force may have fallen to zero, if the oscillations are sufficiently rapid, before the energy, whose velocity of propagation in the conductor is much less than in air owing to the high dielectric constant, has penetrated far into the conductor. This happens at each half-oscillation, and explains the fact that the dissipation of energy is confined to the surface layers when the alternation is sufficiently rapid—a fact which is known as the "skin" effect, and was described on p. 365.

Pressure on Surface due to Incident Wave.—On meeting a plane surface normally, the wave may be transmitted, reflected or absorbed. In general all three processes occur, but in the case of total absorption we may deduce from the conception of the Faraday tubes, which exert a lateral pressure as well as a longitudinal pull, that the wave would exert a pressure upon the absorbent surface. A surface of this kind which absorbs all the radiation falling upon it, is called a perfectly black surface, and

whatever becomes of the radiation which is absorbed, it certainly ceases to be an electromagnetic wave. Thus, if the tubes of induction are destroyed as they meet the surface, we have upon one side of the surface a lateral pressure $\frac{kR^2}{8\pi}$ due to the Faraday tubes (p. 134), which is unbalanced by any tubes on the other side. The pressure is therefore exerted on the surface itself, and since the mean value of R^2 for a whole wave is $\frac{1}{2}R_0^2$, the mean pressure exerted by the Faraday tubes is $\frac{kR_0^2}{16\pi}$. In a similar manner we have a pressure $\frac{\mu\beta_0^2}{16\pi}$ due to the magnetic intensity. But as on p. 424 $kR_0^2=\mu\beta_0^2$, so that the mean total pressure due to both sources is $\frac{kR_0^2}{8\pi}=\frac{\mu\beta_0^2}{8\pi}$, and this is equal to the energy per unit volume of the dielectric through which the wave passes.

This result may be obtained directly from the principle of the conservation of energy. Also, it was predicted by Maxwell as the result of his theory of electromagnetic radiation. He further pointed out, from the identity in magnitude of the velocity $\frac{1}{\sqrt{k_0\mu_0}}$ of plane electromagnetic waves in empty space with that of light, that light consists of electromagnetic waves of very high frequency. The actual existence of electromagnetic waves was not proved during Maxwell's lifetime, but in 1888 Hertz detected them in the neighbourhood of a circuit in which electrical oscillations were occurring.

The existence of a pressure exerted by light falling upon a material surface was demonstrated and measured by Lebedef,[1] who focused a beam of light upon a blackened platinum surface, delicately suspended in a vessel having high vacuum. Care was taken to eliminate disturbances due to variation in temperature, and the result then indicated a pressure due to the incident beam, of the order of magnitude predicted by Maxwell.

A similar measurement was made by Nichols and Hull,[2] an exhaustive set of experiments being carried out. A torsion balance consisting of two polished silver discs suspended by a quartz fibre is situated in an enclosure in which the gas pressure can be varied. Light from an electric arc can be directed upon either disc, and the rotation observed, the intensity of the incident radiation being measured by means of the bolometer. The most important source of error is due to the "radiometer" effect discovered by Crookes, which is of such frequent trouble in deter-

[1] P. Lebedef, *Rapp. Congrès Internat. d. Phys.*, T. 2, p. 133. 1900.
[2] E. F. Nichols and G. F. Hull, *Proc. Amer. Acad.*, **38**, p. 559. 1903.

mining mechanical effects at low gas pressures. The side of the disc upon which the radiation falls, rises in temperature, and the increased velocity of the gas molecules as they rebound from this face results in an excess of pressure upon it, and this effect might be confused with that due to the direct pressure of the radiation. The " radiometer " effect is slowly established, while the radiation pressure is instantaneous, and hence the ballistic observations only are used. The radiation pressure found is in close agreement with the calculated value.

The following list of the results of the determination of the velocity of light affords further proof of the fact that light is an electromagnetic radiation, since the remarkable agreement of the mean value with that for the quantity $\frac{1}{\sqrt{k_0\mu_0}}$ obtained by the methods described in the last chapter can hardly be accidental.

Fizeau	$3{\cdot}150\times10^{10}$	cm. per sec.
Cornu	$3{\cdot}004\times10^{10}$	,,
Foucault	$2{\cdot}980\times10^{10}$	,,
Michelson	$2{\cdot}998\times10^{10}$	,,
Newcomb	$2{\cdot}999\times10^{10}$	,,
Accepted value . . .	$2{\cdot}99790\times10^{10}$	,,

Index of Refraction of Light.—On the wave theory of light it follows that the index of refraction of light on passage from one medium to another (n) is given by,

$$n=\frac{\text{velocity of light in first medium}}{\text{velocity of light in second medium}}$$

Since all transparent media have very nearly the same magnetic permeability, the velocity of a plane electromagnetic wave $\propto\frac{1}{\sqrt{k}}$, and hence if the dielectric constant of the first medium is k_1, and that of the second k_2—

$$n=\sqrt{\frac{k_2}{k_1}}.$$

If the light is passing from air or vacuum to the substance, then $\frac{k_2}{k_1}=k$, the dielectric constant of the medium, taking that of a vacuum as unity, and we have, $n=\sqrt{k}$, or, $k=n^2$.

By the method of discharging a condenser through a ballistic galvanometer, using the gas as dielectric, Klemencic [1] found the

[1] I. Klemencic, *Sitzungsber. Wien. Akad.*, 91 (2), p. 712. 1885.

following values of k; the corresponding values of n^2 are given alongside of them for comparison :—

	k.	n^2.
Air	1·000586	1·0005854
Hydrogen . . .	1·000264	1·0002774
CO_2	1·000984	1·0009088
CO	1·000694	1·0006700

The following values of k are chosen from various sources, and the corresponding values of n^2 placed with them :—

	n^2.	k.	Observer.
Water	1·78	2·79 when $\lambda=1200$	
CS_2	1·99	2·64	Drude, *Zeit. Phys. Chem.*, **23**, 1897
Paraffin . . .	2·02	2·14	Thwing, *Zeit. Phys. Chem.*, **14**, 1894
Sulphur . . .	4·47	2·4	J. J. Thomson, *Proc. Roy. Soc.*, **46**, 1889
Glass (crown) .	2·38	3·24	Gordon, *Phil. Trans.*, **170**, 1879
Glass (light flint) .	2·53	3·01	,, ,, ,,

It will be noticed that in many cases the agreement between n^2 and k is good, but in others, particularly in the case of water, there is an apparent want of agreement. This is probably due to the fact that the effect of absorption has been neglected, and since this may modify the refractive index profoundly, the comparison is only of value in cases where we are certain that absorption of the wave does not exist, as, for example, in the case of the gases. The agreement improves as shorter and shorter waves are taken.

General Case of Wave Propagation.—A more general solution to the equations (iii) (p. 416) and (iv) (p. 418) may be obtained by differentiating (iii) with respect to t, and then eliminating P, Q and R by means of (iv).

Thus from the first of equations (iii)—

$$\mu\frac{d^2\alpha}{dt^2}=\frac{d^2R}{dydt}-\frac{d^2Q}{dzdt}.$$

Now, from the second and third of equations (iv)—

$$-k\frac{d^2Q}{dzdt}=\frac{d^2\alpha}{dz^2}-\frac{d^2\gamma}{dxdz},$$

$$-k\frac{d^2R}{dydt}=\frac{d^2\beta}{dxdy}-\frac{d^2\alpha}{dy^2},$$

$$\therefore\ -k\mu\frac{d^2\alpha}{dt^2}=\frac{d^2\beta}{dxdy}-\frac{d^2\alpha}{dy^2}-\frac{d^2\alpha}{dz^2}+\frac{d^2\gamma}{dxdz},$$

or,

$$k\mu\frac{d^2\alpha}{dt^2}=\frac{d^2\alpha}{dy^2}+\frac{d^2\alpha}{dz^2}-\frac{d^2\beta}{dxdy}-\frac{d^2\gamma}{dxdz}.$$

But differentiating the second equation of (i) (p. 414) with respect to x,

$$\frac{d^2\alpha}{dx^2}+\frac{d^2\beta}{dxdy}+\frac{d^2\gamma}{dxdz}=0,$$

and substituting we get,

$$k\mu\frac{d^2\alpha}{dt^2}=\frac{d^2\alpha}{dx^2}+\frac{d^2\alpha}{dy^2}+\frac{d^2\alpha}{dz^2}.$$

which is usually written, $k\mu\frac{d^2\alpha}{dt^2}=\nabla^2\alpha$. Similar equations may be obtained for β and γ; and further, by differentiating any one of equations (iv) with respect to t and making use of equations (iii), we find that $k\mu\frac{d^2\mathrm{P}}{dt^2}=\nabla^2\mathrm{P}$, with similar equations for Q and R.

A solution of the type

$$\mathrm{P}=f_1(lx+my+nz-vt)+f_2(lx+my+nz+vt)$$

may be found for these equations, in which l, m and n are the direction cosines of the normal to the plane

$$lx+my+nz=p,$$

p being the length of the normal between the origin and the plane, and $v^2=\frac{1}{k\mu}$. Hence the possibility of the propagation of a plane wave in any direction, with velocity $v=\frac{1}{\sqrt{k\mu}}$, which has the value $c=\frac{1}{\sqrt{k_0\mu_0}}$ in empty space (see p. 394).

This general form is of great use in the theory of light, and for a further development of it the student is referred to works on optics.

Oscillatory Discharge.—We have already seen that the equation for the electromotive forces occurring in a conductor which has capacity and inductance, leads to the conclusion that any change in the electrical condition of the conductor is accompanied by oscillations when $\frac{\mathrm{L}}{\mathrm{C}}>\frac{\mathrm{R}^2}{4}$ (p. 337), and that the time of one complete oscillation is $\frac{2\pi}{\sqrt{\frac{1}{\mathrm{LC}}-\frac{\mathrm{R}^2}{4\mathrm{L}^2}}}$ (p. 338), which reduces to $2\pi\sqrt{\mathrm{LC}}$ when R is small. We are now in a position to interpret this process in terms of the Faraday tubes and their motion in the surrounding dielectric. If the circuit consist of two conductors

A and B separated by a small air gap, and A have a negative and B a positive charge, the distribution of the electrical field will be somewhat as shown in Fig. 355 (i), the lines indicating Faraday tubes of induction, the tubes on one side only being drawn. On increasing the charges upon A and B, the difference of potential between them will rise until a certain limit is reached depending upon the length of the gap, the nature of the electrodes, and the pressure of the air. As soon as this limit is reached, the gap suddenly becomes conducting, and the negative ends of the tubes situated upon A move downwards, those upon B moving upwards, both constituting a positive electric current flowing upwards.

If the tubes did not possess inertia they would all in turn collapse in the gap, those at the outside shrinking to fill the space previously occupied by those which have disappeared, and there would be no oscillation. But the Faraday tubes in motion are

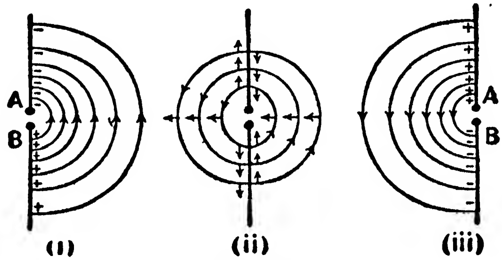


FIG. 355.

accompanied by a magnetic field, and, as was shown by Sir J. J. Thomson (p. 422), the magnetic field at right angles to their length and to their direction of motion being given by $H=4\pi Dv \sin\theta$, where θ is the angle between the tube and its direction of motion.

Now, the magnetic field H is associated with an amount of energy equal to $\frac{\mu H^2}{8\pi}$ per unit volume,

that is, $$\frac{16\pi^2 D^2 v^2 \mu \sin^2\theta}{8\pi} = 2\pi\mu D^2 v^2 \sin^2\theta$$

and this is consequently the energy of the Faraday tubes due to their velocity v. By analogy with the expression $\frac{1}{2}mv^2$ for the kinetic energy of a moving body of mass m, we may imagine the tubes to be endowed with mass $4\pi\mu D^2 \sin^2\theta$ per unit volume, and their momentum would therefore be $4\pi\mu D^2 v \sin^2\theta$ when moving with velocity v at an angle θ to their length. Thus the mass per unit volume is zero when moving in their own direction, for $\theta=0$, and $4\pi\mu D^2$ when moving at right angles to their length.

The conception of the Faraday tube has been developed for the purpose of explaining the attractions and repulsions between electrical charges, the tension along the tube tending to pull together the opposite charges at the ends, and the lateral pressures between contiguous tubes pushing like charges apart. These lateral pressures, however, can only exist when the neighbouring tubes have the same direction. To account completely for the phenomena, we must imagine that oppositely directed tubes attract each other and may even coalesce on meeting. Thus if two pairs of charges AB and CD be placed at a distance apart upon two conductors as in Fig. 356 (i), although a few Faraday tubes may exist between A and C, and B and D respectively, yet the greater number will exist between A and B, and C and D. Let the charges A and C approach and neutralise each other, as will B and D. The process may be looked at from two points of view. Either the few tubes between A and C pull the charges together, shrinking in the process and eventually disappearing, as also do those between B and D, in which case the tubes from A to B, and from D to C eventually coincide in position and have a zero resultant; or we may consider that the two sets of tubes, being oppositely directed, attract each other. When the approaching tubes meet they coalesce and break up into tubes joining AC and BD respectively, as at e and f (Fig. 356 (ii)); these then shrink and eventually disappear. The result is the same on either supposition; the charges have met and neutralised each other, there having been no metallic connection at any time between the two conductors.

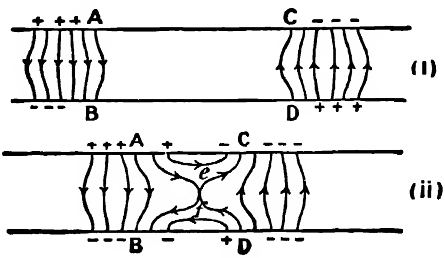


FIG. 356.

If, however, the charges approach each other with the velocity of propagation of an electromagnetic wave, the state of affairs is different. Each moving set of tubes has its accompanying magnetic field, and since the tubes are oppositely directed and their velocities are also opposite, their magnetic fields coincide in direction, being directed through the plane of the paper from front to back in Fig. 356. The energy still exists in the form of the magnetic field even at the instant that the charges and Faraday tubes are neutralising each other. Thus each still has its own momentum, and the two will cross and then recede from each other with undiminished velocity.

If two sets of Faraday tubes having the same direction are

approaching each other, their magnetic fields as the two waves approach will be oppositely directed and give zero magnetic field at the instant of coincidence. But the electrical charges and displacements are in this case doubled at this instant owing to the coinciding of the two sets, and the energy in this case persists in the form of the electrical field. This gives rise to two equal waves which recede from each other with equal velocities, and we may to all intents and purposes say that the two waves have crossed each other and passed on, each unaffected by the other. Hence we may consider that when travelling, the Faraday tubes, having inertia on account of their accompanying magnetic fields, can cross each other, each passing on unaffected by the other.

Returning to the consideration of Fig. 355, we see that when the gap becomes conducting, the lateral pressure on the tubes near the gap due to those lying outside, will cause them to travel inwards, the ends lying upon A travelling downwards, and those upon B upwards, but every part of the tube, since it arrives at the gap with velocity, and therefore momentum, will continue to travel onwards, and the ends will cross each other at the gap, the positive end will travel up A, and the negative end down B, the tube meanwhile spreading out on the other side of AB. The state of affairs when half the tubes have crossed the gap is shown in Fig. 355 (ii), and all the tubes there drawn are at this instant travelling from right to left. The process will continue until all the tubes have crossed. They will come to rest as in Fig. 355 (iii) when their momentum has been reduced to zero by the stresses in the tubes, which will now tend to drive them to the right. A is now positively charged and B negatively; in fact, the current has flowed from B to A until this reversal has been effected. The current then ceases, and is ready to begin the reverse flow if the gap is still conducting. For the purpose of clearness only half the field has been drawn in the diagram, but it will readily be seen that the actual state of the field will be obtained by revolving the figure about AB. In (ii) the tubes which at the start were on opposite sides of AB are now half on each side, and the value of D, or the resultant number of tubes per unit area, is zero, since the tubes have reversed their direction on passing the gap. The resultant electrical charge and displacement at this instant are everywhere zero; but it must be remembered that the two sets of tubes at any point have opposite velocities as well as directions, and therefore the magnetic fields corresponding to them are coincident in direction, and their resultant is obtained by adding the two together. They are therefore circles surrounding AB, and the energy of the charge is now in the form of the magnetic field. The direction of the magnetic field is that due to the current flowing upwards in AB.

As the current surges backwards and forwards between A and B the energy of the system alternates between the electrostatic and magnetic forms, and if the former be compared to potential energy the latter is kinetic, and the energy of the system, like that of a mechanical vibrating system, alternates between the static and the kinetic forms.

The presence of the air gap is not essential to the above discussion ; if positive and negative charges be simultaneously given to two ends of a conductor, or even if only one charge be given at a point of it, the redistribution of the charge will be accompanied by surgings backwards and forwards, which will be eventually damped out on account of the resistance of the conductor itself to the passage of the current through it, and the energy of the vibration will gradually be dissipated in the form of heat in the conductor.

Rapid Oscillations and Radiation.—Provided that the oscillations are sufficiently slow, the energy is entirely transformed into heat in the conductor, none of it passing permanently into the surrounding dielectric, but with rapid oscillations this is no longer true.

Starting with a shorter conductor, which will of course have less capacity and inductance, the lines on one side of it at the

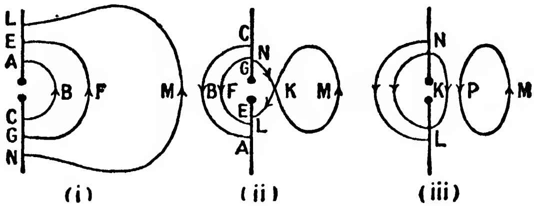


FIG. 357.

moment at which the gap begins to be conducting will be represented by Fig. 357 (i). When the discharge begins, the ends of the tubes near the gap will cross ; the tubes become reversed, exactly as in the previous case. A tube such as ABC (Fig. 357 (i)) will be reversed on reaching the gap, and will spread out upon the other side, on account of its own inertia and the pressure of the tubes behind it, whereas the ends L and N of a tube such as LMN will reach the gap before the equatorial part M, and will cross, as at K (Fig. 357 (ii)). A stage in the process will be reached when the branches LKM and NKM at the point of intersection, are moving parallel to their own directions, and they have then no momentum to carry them past each other and coalescence occurs, the tube separating into a closed loop PM and a half loop LKN (iii) with ends upon the conductor, which

continues to grow, owing to the momentum of the parts of the tube at L and N. This part of the tube will cease to grow when its own tension and the pressure of the neighbouring tubes have exhausted its momentum.

There will be some limiting tube EFG, such that those within it cross to the other side of the conductor, while those lying outside it will form loops and will not pass the conductor.

We are now in a position to understand the processes going on when radiation occurs.

Starting with the Faraday tubes as shown in Fig. 358 (i), those

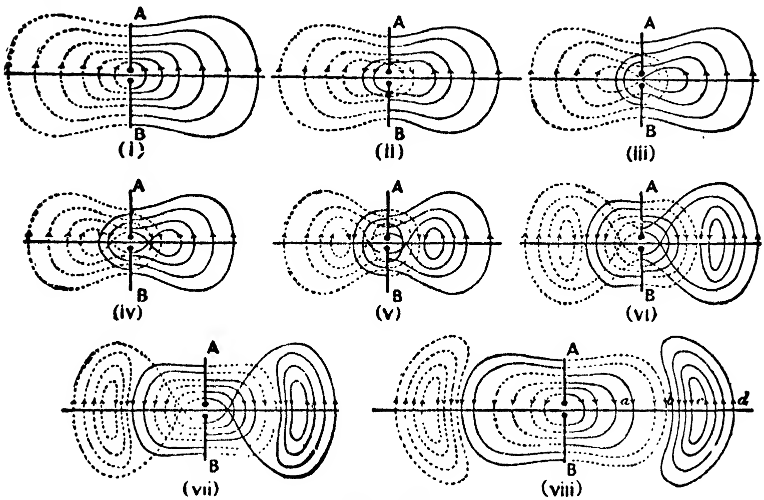


FIG. 358.

to the right of AB at the start are drawn in full line, and those to the left dotted. A has a negative and B a positive charge at the moment at which the gap becomes conducting. In (ii) the first two tubes have crossed the gap and will afterwards continue to expand until the first half-oscillation is complete. At (iii) the ends of the third tube have met at the gap before the equatorial part reaches it, and the ends cross, forming a loop, as at (iv). At (v) the third line has broken into two parts forming the closed loop, and the fourth line is in the act of breaking. In (vi), (vii), and (viii) the process is continued, and the remaining tubes form closed loops.

As each tube breaks, the tension in each part of it causes a pull which brings the closed part into the space between the two tubes which have crossed from the other side and the loop immediately inside it, and the remainder which has its ends upon the conductors is similarly pulled to the inside of the adjacent tube. Whether the tubes are supposed to cross each other, or whether

they reach their final state by frequent coalescences between neighbouring tubes with reseparation in their new form, is immaterial, as we saw on p. 433.

The closed loops, immediately after their formation, are pushed outwards by the expanding tubes behind them, and when the condition represented in (viii) is reached, the loops are travelling outwards and will now continue to travel with the velocity $\frac{1}{\sqrt{k_0\mu_0}}$. In the next half-oscillation a second set of loops of reverse sign will be pushed outwards; the whole set will, as they spread outwards, become cylindrical sheets which on expanding, continually approximate to planes of electrical displacement normal to the line joining them to the middle of the oscillator. Those parts of the loops that travel away from the equatorial plane need not here concern us.

Relative Phases of Electrical and Magnetic Fields.—We have already seen (p. 423) that in an electromagnetic wave, the magnetic field is in phase with the electric displacement, that is they both reach their maximum values and their zero values simultaneously. Thus at points such as b (Fig. 358 (viii)) the magnetic field and the electric displacement have their maximum values, at c they are both zero, and at d they again have maximum values, but of opposite sign to those at b. At the oscillator the magnetic field and the electrical displacement are 90° apart in phase; that is, one has its maximum value when the other is zero, as we saw on p. 434. Hence between the oscillator and the point b one of them has been displaced 90° in phase relatively to the other. It will be seen that the phase of the electrical displacement is the same at b as at the oscillator, and therefore in calculating the phase of the electromagnetic wave at a distant point at any instant, the distance of the point from b must be used, and not the distance from the oscillator. This fact bears a remarkable analogy to the quarter wave-length discrepancy that occurs when calculating the phase of the light vibration due to a plane wave of light, at a point in advance of the wave-front. On splitting up the wave front into Fresnel's zones, it is found that the resultant effect of the wave is that due to half the first Fresnel's zone at the pole of the wave, but to get the correct phase, the wave-front must be imagined to be *displaced forward by a quarter of a wave-length.* The analogy between this case and that of the electromagnetic oscillator was pointed out by Professor F. T. Trouton,[1] and he calculated from Hertz's equations that the phases of the magnetic field and electric displacement at a distance from the oscillator are correct, if the distance

[1] F. T. Trouton, *Phil. Mag.* (Ser. 5), **29**, p. 268. 1890.

be measured from the point b and not from the oscillator, when the distance from the oscillator to b is $\frac{\lambda}{4 \cdot 4}$, λ being the wave-length of the disturbance at a distance from the oscillator.

We may get some idea of the reason for this change in the relative phases of the field and displacement by drawing a curve for each, for the positions a, b, c, d in Fig. 358 (viii). At the oscillator the displacement has reached its negative maximum value, and is for an instant stationary, and from O to some point between a and b the Faraday tubes are on the point of beginning to return for the second half-oscillation, but at b they are moving outwards. At c there is zero, and at d the maximum positive displacement. The full-line curve D in Fig. 359 represents the distribution of displacement. The magnetic field is zero where the tubes are at rest; but increases in value between a and b, reaching at b a maximum. At c it is zero, and at d again a maximum. Its distribution is represented by the full-line curve H.

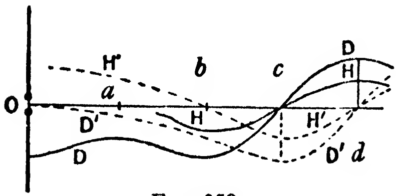


FIG. 359.

A quarter of a period later the condition is represented by the dotted curves D′ and H′ in Fig. 359. Near the oscillator there is a magnetic field due to the motion of the Faraday tubes towards it, while at b the field is zero. The maximum previously at b has reached c, and is in phase with the displacement. The succeeding oscillations then produce a train of electromagnetic waves of the ordinary type travelling outwards from b.

Reflection of Plane Waves.—In the case of the oscillator, we have seen that some of the Faraday tubes of the surrounding field, instead of passing the conductor when they collapse upon it, are reflected. A similar reflection occurs when plane electromagnetic waves meet a conducting surface. For simplicity we will consider a plane electromagnetic wave of the type described on p. 420, falling normally upon a perfectly conducting surface.

Let the equation of the electric intensity be

$$R = R_0 \sin 2\pi\left(\frac{x}{\lambda} + \frac{t}{T}\right).$$

From p. 420 we see that this is a wave travelling towards the origin with velocity $v = \frac{\lambda}{T}$.

The magnetic field is then obtained from the relation $\mu\frac{d\beta}{dt} = -\frac{dR}{dx}$ (p. 418); *i.e.*—

$$\frac{d\beta}{dt} = -\frac{2\pi R_0}{\mu\lambda}\cos 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right),$$

and,

$$\beta = -\frac{TR_0}{\mu\lambda}\sin 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right)$$

$$= -\beta_0 \sin 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right),$$

where $\beta_0 = \frac{T}{\mu\lambda}R_0$.

The curves for R and β are drawn in Fig. 360 for the instant when $t=\frac{T}{4}$, and the wave is completely represented by the motion of the curves A and C from right to left with velocity $v=\frac{1}{\sqrt{k\mu}}$.

If then YOZ be a perfectly conducting plane, the electric intensity in this plane must always be zero. Hence in the plane itself

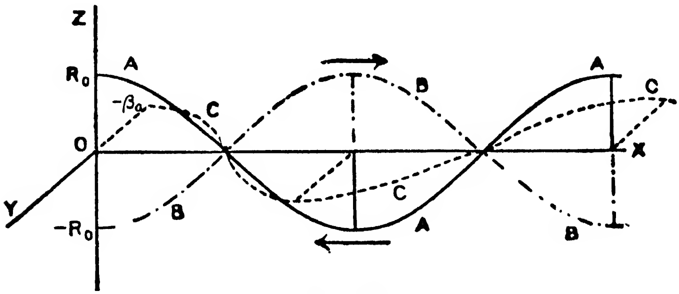


FIG. 360.

some electric intensity is brought into play which is equal and opposite to R at every instant, for otherwise there would be a resultant intensity differing from zero. The electric intensity at the plane is obtained by putting $x=0$ in the equation for R, whence $R=R_0 \sin 2\pi\frac{t}{T}$, and this opposite intensity created in the conducting plane is therefore $R=-R_0 \sin 2\pi\frac{t}{T}$.

A harmonic disturbance such as this gives rise to two sets of waves, one travelling to right and one to left from the point. The one to the left is $R=-R_0 \sin 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right)$, and is in the direction of the incident wave, and by comparing the equations for the two, we see that the two waves entirely cancel each other out, and there is no effect beyond the conducting plane. This is just as we should expect, for a perfect conductor is opaque to

electromagnetic radiation. The other wave, travelling to the right, has the equation $R=R_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)$, for this reduces to $R=-R_0 \sin 2\pi\frac{t}{T}$ at the plane, and has the form of a wave travelling to the right.

The instantaneous intensity R at the conducting plane will give rise to a current in it, whose accompanying magnetic field is β, and is in the direction given by Fig. 351. It must be remembered that in Figs. 350 and 360, the direction of the axis OY is arbitrarily chosen, but whatever direction for the axis is chosen, the actual positive direction of the magnetic field must always be in accordance with that in Fig. 351.

In a similar manner, the magnetic field β, meeting the plane, gives rise to an electric intensity R in it. For a strip of unit width parallel to OZ will, for every unit length of the strip, be cut by magnetic induction at the rate $\mu\beta v$ units per second, and this is equal to the electromotive force in it.

But,
$$\beta=-\frac{T}{\mu}\cdot\frac{R_0}{\lambda}\sin 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right)$$
$$=-\frac{T}{\mu\lambda}\cdot R.$$

Now, $\frac{\lambda}{T}=v$, and E.M.F. $=\mu\beta v$
$$=\frac{T}{\mu\lambda}\cdot v\mu R=R.$$

Another way of looking at the process of reflection is to consider that as the Faraday tubes arrive at the conducting plane, their ends on reaching the plane travel together, and the tube becomes reversed on account of its inertia. The magnetic field is not reversed, so that by Poynting's Theorem (p. 425) the direction of propagation of energy is reversed; that is, the wave now travels away from the plane. The reflected wave is represented at the instant $t=\frac{T}{4}$ by the curves C and B (Fig. 360), the curve C at the given instant being the same for both incident and reflected waves, for the accompanying magnetic field β to the wave $R=R_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)$ is, as we have seen on p. 421—
$$\beta=\beta_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)$$

and is in phase with the magnetic field of the incident wave at

the plane, since here $x=0$. For this reflected wave which is travelling to the right, β and R have the same sign, and it may be noticed that for the incident wave which is travelling to the left, β and R have opposite signs; this also gives us a means of determining the direction of propagation of the wave, and is in accordance with Poynting's theorem. It should be remembered that Fig. 360 is drawn for the epoch $t=\frac{T}{4}$.

If the conducting plane be divided into strips by a number of parallel non-conducting lines, the reflection of the wave is unaffected when these strips are parallel to the direction of R, since the conductivity in the direction of R is unchanged, but when at right angles to R, the wave as it meets the plane cannot produce any current. In fact, it is now non-conducting in the direction of the electromotive force, if the strips are sufficiently narrow, and reflection will not occur. We shall see later that Hertz used a metallic grating, and found that when the metallic strips are parallel to the electric displacement, reflection occurred, but not when they are at right angles to it; the waves in this case are transmitted. It will be seen that such a grating behaves towards an electromagnetic wave exactly as a Nicol's prism behaves towards light. Two such gratings may be used as polariser and analyser respectively.

Stationary Oscillation.—We should expect, from analogy with the case of sound waves and waves in stretched strings, that the two waves, the incident and the reflected ones described above, would combine to produce a condition of steady oscillation.

If we find the resultant electric intensity due to both incident wave $R=R_0 \sin 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right)$, and reflected wave $R=R_0 \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)$, we have—

$$R_1=R_0\left\{\sin 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right)+\sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)\right\}$$

$$=2R_0 \sin 2\pi\frac{x}{\lambda} \cos 2\pi\frac{t}{T}$$

This represents a state of steady oscillation in electrical condition, for at any given point, x is constant and the oscillation is harmonic of the type $R_1=A \cos 2\pi\frac{t}{T}$, and it will be seen that the phase is now independent of x. The amplitude A itself varies with x according to the equation $A=2R_0 \sin 2\pi\frac{x}{\lambda}$. It is therefore zero at the reflecting surface, and reaches its maximum value,

$2R_0$, at a distance $\frac{\lambda}{4}$ from the surface. The successive values of R during half an oscillation are indicated by curves 1, 2, 3, 4, 5 in Fig. 361 (i).

The value of β at any point is similarly obtained by adding the values for the incident and reflected waves—

$$\beta_1 = -\beta_0\left\{\sin 2\pi\left(\frac{x}{\lambda}+\frac{t}{T}\right) - \sin 2\pi\left(\frac{x}{\lambda}-\frac{t}{T}\right)\right\}$$

$$= -2\beta_0 \cos 2\pi\frac{x}{\lambda} \sin 2\pi\frac{t}{T}.$$

This is also the equation of a stationary oscillation, whose amplitude is $2\beta_0$, where $x=0$, that is at the reflecting surface, and also

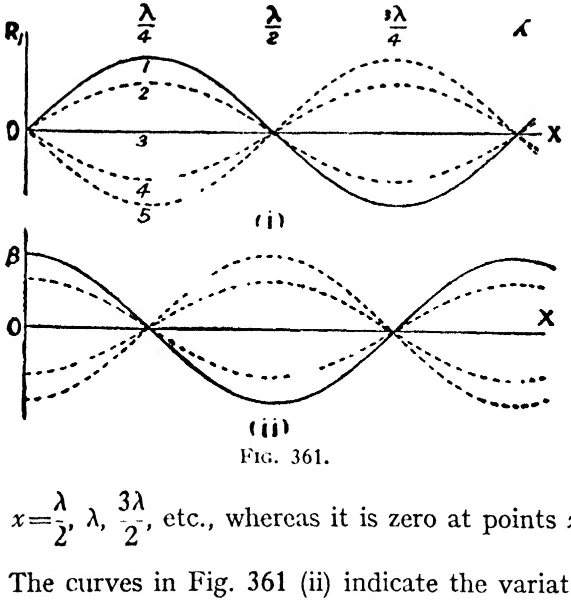


FIG. 361.

at points $x=\frac{\lambda}{2}$, λ, $\frac{3\lambda}{2}$, etc., whereas it is zero at points $x=\frac{\lambda}{4}$, $\frac{3\lambda}{4}$, $\frac{5\lambda}{4}$, etc. The curves in Fig. 361 (ii) indicate the variation in β_1 during a half-oscillation; but it must be remembered that the direction of β_1 is at right angles to that of R_1; the ordinates in the diagrams merely represent to scale the respective values of β_1. By comparing (i) and (ii) we see that for points where the fluctuation in R is a maximum, the fluctuation in β is zero, and *vice versa*.

Parallel Wires.—On connecting one end of each of a pair of parallel wires to the terminals of a condenser in which an oscillation in electrical condition is occurring (see p. 436), electromagnetic waves are produced which travel down the wires. The waves are not in this case plane, the ends of the Faraday tubes

being confined to the wires in a manner similar to that for the parallel planes described on p. 422. Although the actual shape of the wave is not easily determined, the process of propagation is like that which we have been considering, and if the far ends of the wires are joined together, reflection will occur there. If the ends are not joined by a conductor, a discharge may occur with production of brushes or sparks, the effect being to produce a reflected wave, although in this case some of the energy of the wave is lost at the point of reflection. Sir Oliver Lodge [1] has shown, by means of two such parallel wires with their near ends connected respectively to the inner and outer coatings of a Leyden jar, that, taking the length of the wires when maximum spark occurs at the free ends to be half the wave-length of the electromagnetic oscillation, the velocity of propagation obtained by multiplying this wave-length by the frequency of oscillation of the jar as calculated from its dimensions, is about that of light-waves.

The arrangement of parallel wires was afterwards improved by Lecher, and is described on p. 459.

Velocity of Propagation.—The velocity of propagation of an electromagnetic disturbance has been directly determined by Blondlot,[2] who used two pairs of cylindrical condensers in parallel, one pair being short circuited by an air-gap, and the other discharging through the same air-gap, but the charge having to pass on the way through two long wires, PLC and QMD (Fig. 362). The insides of two glass cylinders are completely covered with tinfoil, and on the outsides are two pairs of rings of tinfoil, AB and CD. These are joined by moistened threads of high resistance shown by dotted lines in the diagram. When a spark passes at S, A and B immediately discharge, giving a spark between the points P and Q. C and D also give a spark, but the charges have to pass through the long wires CLP and DMQ, each of length 1029 metres, the spark therefore occurring later than that due to A and B. The interval between the sparks is obtained by measuring the distances between their images upon a photographic plate, produced by a rotating concave mirror, and is the time taken for the electromagnetic wave to travel a distance of 1029 metres along the wire. In this way the velocity of propagation was found to be $2{\cdot}96 \times 10^{10}$ cm. per

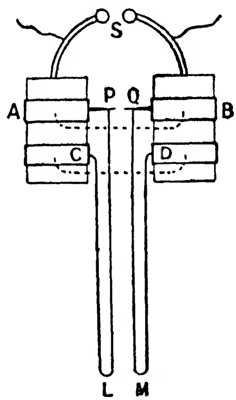


FIG. 362.

[1] O. J. Lodge, *Phil. Mag.* (Ser. 5), 26, p. 217. 1888.
[2] R. Blondlot, *Comptes Rendus*, 117, p. 543. 1893.

second, and a second set of experiments with wires 1821·4 metres long gave $2{\cdot}98 \times 10^{10}$ cm. per second.

Hertz's Experiments.—The prediction by Lord Kelvin in 1853, from mathematical reasoning, that the discharge of a Leyden jar would under certain conditions be oscillatory (p. 338), was followed in 1857 by the demonstration of such oscillations by Feddersen on examining the spark in a rotating mirror. In 1865 Maxwell published his theory of electromagnetic radiation, but it was not until 1888 that Professor H. Hertz [1] proved the existence of such radiations in the space surrounding a Leyden jar in which electrical oscillations were occurring.

The variety of oscillators used for the production of electromagnetic waves is very great, but one of the earliest forms, used by Hertz himself, consists of two square sheets of metal, having

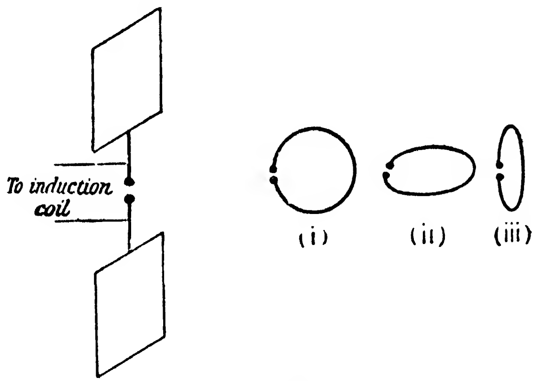


FIG. 363.

sides 40 cm. in length, placed about 60 cm. apart, and having two gilt and highly polished balls 2 or 3 cm. apart and connected to the plates by light metallic rods (Fig. 363). The plates are connected to the terminals of an induction coil, and every time the difference of potential between the balls reaches a sufficient value to render the air in the gap conducting, an oscillatory discharge occurs, with radiation of electromagnetic waves. The balls must be kept highly polished, or the beginning of the discharge will not be sufficiently abrupt for the production of radiation.

The period of oscillation (T) of this apparatus is about $1{\cdot}8 \times 10^{-8}$ second, and the velocity of propagation (c) being 3×10^{10} cm. per second, we see that the wave-length (λ) given by $\lambda = cT$, is $5{\cdot}4 \times 10^2$ cm., or 5·4 metres.

In order to detect the radiation, Hertz employed a circle of wire 35 cm. in radius, with a gap at one point, and here sparks

[1] H. Hertz, *Nature*, **39**, pp. 402, 450, 547 (1889); and *Wied. Ann.*, **1**, 1889.

occur when the radiation is falling upon the detector in a suitable manner. The detector itself has a proper period of its own for oscillation of charge between the knobs. If opposite charges are given to the two knobs of the detector, the Faraday tubes will contract, the two ends travelling round the circular conductor to meet each other. They will therefore, on account of their inertia, cross at the opposite end of a diameter to the gap, and then grow in the reverse direction until the charges on the ends have become completely reversed. If the period of oscillation for the detector is equal to that of the electromagnetic waves falling upon it, the Faraday tubes of the next half-wave will, on reaching the gap, cause the knobs to be oppositely charged ; or we may use the alternative explanation that the magnetic field at right angles to the gap, as it travels across it, will produce an electromotive force between the knobs. Then the reversal of the charge is just completed at the instant when the next half-wave of radiation arrives, and this will assist the reversed charging of the knobs, and the next half-oscillation will be more violent than the last. Or, to borrow an expression from acoustics, the detector resounds to the waves, or resonance occurs. This type of detector is therefore called a resonator.

In using a given detector, the position of the knobs is adjusted until maximum sparking between them occurs. The period of its proper oscillation is then the same as that of the incident wave.

The orientation of the resonator is of importance, since the electromagnetic wave is polarised, that is, the electrical displacement is, for points in the equatorial plane, in one direction only —parallel to the gap of the oscillator. The gap of the resonator must be parallel to this direction, and thus the resonator will detect the oscillations when in positions (i) and (iii), but not when in position (ii).

Refraction of Waves.—Using a reflector consisting of a metal sheet bent into a parabolic form with the oscillator in the focal line, the waves may be restricted to a comparatively narrow beam in which the wave front is plane. A similar reflector with resonator in its focal plane is used as a detector. The beams consist of polarised waves, as may be shown by placing the reflectors first with their focal lines parallel, and then with them at right angles. In the former case sparking of the resonator occurs, but not in the latter. With such an arrangement Professor Trouton repeated many of the ordinary optical experiments, using prisms of pitch or paraffin wax, and determining the index of refraction for these materials. Using a paraffin wall 3 feet in thickness he showed [1] that reflection takes place for all angles of incidence, provided that the electric displacement in

[1] F. T. Trouton, *Nature*, 39, p. 391. 1889.

the beam is perpendicular to the plane of incidence, but when the electric displacement is in this plane, there is some angle of incidence for which reflection does not occur. Hence the electric displacement is perpendicular to the plane of polarisation. according to its optical definition.

Determination of Wave-Length by Stationary Oscillations.—By means of apparatus of the kind described on p. 444, Hertz demonstrated the existence of stationary oscillations of the type described on p. 442. The radiation was allowed to fall on a large plane sheet of zinc, at which reflection occurs, and the incident and reflected waves together form a steady vibration. It was found that as the resonator is moved outwards from the sheet, it goes through a series of maximum and minimum excitation, the maxima corresponding to points at distances $\frac{\lambda}{4}, \frac{3\lambda}{4}$, etc., from the sheet, and the minima to points $\frac{\lambda}{2}, \lambda, \frac{3\lambda}{2}$, etc. (Fig. 361 (i)).

In this way the wave-length of the emitted waves could be determined, and knowing the frequency of oscillation, the velocity of propagation was found. With a periodic time of about $1{\cdot}8\times10^{-8}$ seconds, the minima are about 2·7 metres apart. This, being half the wave-length, gives a velocity of propagation of about $3{\cdot}0\times10^{10}$ cm. per second.

The interpretation of these experiments is not quite satisfactory. It was pointed out by Sarasin and De la Rive [1] that the distance between the nodes in Hertz's experiments depends rather upon the time of natural oscillation of the detector than upon that of the oscillator. The oscillator being of the "open" type, the oscillations are highly damped, only a very few complete waves being emitted at each discharge. Hence no interference between the incident and reflected waves could be expected; but these waves serve to start oscillations in the detector, which, being of the "closed" type, will emit waves very slightly damped, and if the distance from the detector to the reflecting wall be an odd number of quarter wave-lengths the reflected wave will be in the right phase to reinforce the vibration. The distance between successive points of maximum disturbance or of minimum disturbance (nodes) will therefore be $\frac{\lambda}{2}$. When the detector is turned to the oscillator, the effect will be as found by Hertz.

Various Oscillators.—Many other forms of oscillator have been used. Sir Oliver Lodge [2] tuned two Leyden jars to the same

[1] Sarasin and De la Rive, *Comptes Rendus*, **112**, p. 658. 1891.
[2] O. J. Lodge, *Nature*, **41**, p. 368. 1890.

period of oscillation by altering the position of the slider S, which makes contact with the parallel wires connected respectively to the inner and outer coatings of the Leyden jar B (Fig. 364). When sparks occur between the knobs of the jar A the electromagnetic waves emitted set up surgings of the charge in B, which, since B is short-circuited by means of S, will grow until the alternating difference of potential between the coatings is sufficient to cause minute sparks to occur at the short air gap C.

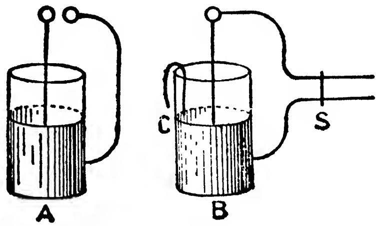


Fig. 364.

Lodge[1] has also obtained oscillations in a single metallic sphere 5 cm. in diameter; they may, however, be much more easily obtained in the case of larger spheres. Minute sparks may be obtained from a similar sphere several yards away, on bringing a conductor into contact with it at the ends of a suitable diameter.

Signalling by Electromagnetic Waves.—The work of Hertz and Lodge has been followed by a number of applications of the principles of electromagnetic waves to signalling, or, as it is commonly called, Wireless Telegraphy or Telephony. Amongst the great number of the principal forms in which the waves have been employed, we shall only trace the history, the student being referred for a comprehensive treatment to the various works on this subject.

By means of the induction coil, a series of discharges takes place between the knobs S (Fig. 365), and at each discharge, oscillations occur in the circuit consisting of the condenser C and one winding of the transformer T. The other winding of the transformer is in series with the long vertical conductor A, called the antenna or aerial, in which it produces an oscillating electromotive force. When the natural period of oscillation for the aerial is the same as that of the condenser circuit, the amplitude of the oscillation in it will be very great. Hence the contact P is adjusted in position until the inductance in series with the aerial is such that the two circuits are in tune. The antenna A is then the seat of radiation of a type similar to that from a Hertzian

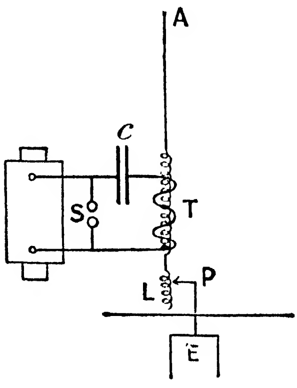


Fig. 365.

[1] *Ibid.*, p. 462. 1890.

oscillator, but since its lower end is earthed, only the upper half of the wave of Fig. 358 is produced. A series of such waves is given in Fig. 366.

At the receiving station, the electromagnetic waves meet a similar antenna, and electric surgings in them are set up.

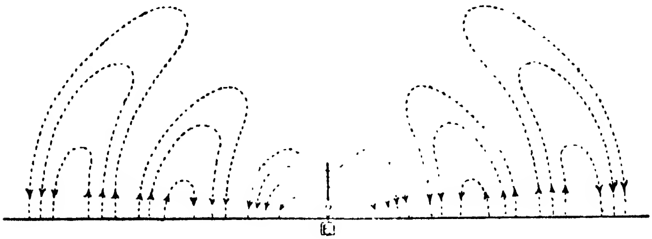


FIG. 366.

Radiation from an Aerial in the Form of a Wire connecting the Two Plates of a Condenser.—The aerial of Fig. 366 is not the most efficient form of radiator, for the current at the top is obviously zero and increases to a maximum at the bottom. For this reason the practice is now adopted of placing a capacity at the top, in the form of a cross wire or network of cross wires. If A (Fig. 367) is such a capacity and the wire of vertical length h connects A and the earth B, the oscillatory current will flow between A and B, and if the capacity of A is sufficiently great, the current will be constant throughout h.

Let the current i consist of a movement of charge q per centimetre with uniform velocity v, all measured in absolute electromagnetic units. Then $i=qv$. Let CD be an electric line of force, which is also moving, at right angles to itself, with velocity v. At any point at distance x from h, let the electric intensity be E,

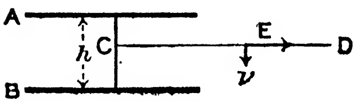


FIG. 367.

then $E=\dfrac{qf(x)}{k}$, where $f(x)$ is a geometrical factor connecting q and E, and involving amongst other things the shape and size of the aerial.

$$\text{Electric displacement } D=\frac{kE}{4\pi} \qquad \text{(p. 128)}$$

$$=\frac{kqf(x)}{4\pi k}=\frac{qf(x)}{4\pi}.$$

But the motion of the electric field is the magnetic field H (p. 422), whose value is given by—

$$H=4\pi \,.\, D \,.\, v$$
$$=qvf(x)=if(x).$$

If x is great compared with h,

$$H=\frac{ih}{x^2} \qquad \text{(p. 48)}$$

$$\therefore f(x)=\frac{h}{x^2}.$$

The deduction of the radiation which follows when the current changes is suggested by the treatment of the radiation of an accelerated charge by Sir J. J. Thomson.[1]

Let the current vary, that is, the charge q have an acceleration (a), positive or negative, for a very short time δt. In this short interval the velocity changes by amount $\delta v=a\,.\,\delta t$. The relative velocity of the line LM (Fig. 368) with respect to the old velocity is δv, and in the interval of time t, large compared with δt, the separation of the portions LM and NG occurring after the acceleration at CD is $FL=NP=t\,.\,\delta v$. The pulse MN which changes the old line to the new, travels outwards with the velocity of light $\frac{1}{\sqrt{k_0\mu_0}}=c$, and at the bend there is a component PN of the electric field, perpendicular to the direction of the line. The shape of the pulse MN is immaterial. Now the separation of the two positions of the line takes place in the time the pulse travels over the distance $MP=c\,.\,\delta t$,

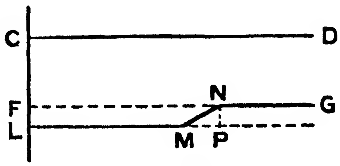


FIG. 368.

$$\therefore \frac{NP}{MP}=\frac{t\,.\,\delta v}{c\,.\,\delta t}$$

and if MP represents the electric displacement D at distance x from the oscillator, $D=\frac{qf(x)}{4\pi}$,

$$\therefore \text{displacement } NP=\frac{t\,.\,\delta v}{c\,.\,\delta t}\,.\,\frac{qf(x)}{4\pi}.$$

Now, $x=LM=ct$, the distance the pulse travels in time t, and $\frac{\delta v}{\delta t}=a$,

$$\therefore \text{displacement } NP=\frac{aqx}{4\pi c^2}f(x).$$

The displacement NP travelling with velocity c is equivalent to a magnetic field $H=4\pi(NP)c$ (p. 422),

$$\therefore H=\frac{aqx}{c}f(x).$$

[1] J. J. Thomson, *Phil. Mag.*, 45, p. 172. 1898.

Again, the current in the aerial is $i=qv$, and in the case where this is oscillating, $i=i_0 \sin pt$,

$$\therefore\ qv=i_0 \sin pt,$$

and
$$q\frac{dv}{dt}=pi_0 \cos pt.$$

But $\dfrac{dv}{dt}=a$,

$$\therefore\ \mathrm{H}=\frac{xpi_0 f(x) \cos pt}{c}.$$

Remembering that $f(x)=\dfrac{h}{x^2}$,

$$\mathrm{H}=\frac{hi_0 p \cos pt}{xc},$$

a result which is in accord with other methods of treatment. If time be reckoned from the moment at which the disturbance leaves the aerial,

$$\mathrm{H}=\frac{hi_0 p \cos 2\pi\left(\dfrac{x}{\lambda}-\dfrac{t}{\tau}\right)}{xc}$$

and the corresponding displacement may be found as on p. 422.

It should be noted that the magnetic field $\dfrac{i_0 h}{x^2}$ is in the same direction as the above, but does not correspond to radiation from the aerial because the corresponding electric displacement is radial, whereas the magnetic field $\dfrac{hi_0 p \cos a\pi\left(\dfrac{x}{\lambda}-\dfrac{t}{\tau}\right)}{xc}$ is radiated, owing to the fact that its corresponding electric displacement is parallel to the aerial, and Poynting's theorem (p. 425) then indicates a radial velocity.

Detection of Electromagnetic Radiation.—In order to detect the current produced in the receiving aerial, the *coherer*, the principle of which was discovered by Sir Oliver Lodge, was used in the early experiments. Lodge [1] found that when electrical oscillations occurred between two metallic surfaces in poor contact, the resistance of the contact at once fell to a very small amount, but was immediately restored to its original value by any mechanical disturbance, such as tapping. Branly used for the same purpose a tube containing metallic filings, and this was again improved by Marconi, who used a mixture of nickel and silver filings (95% nickel) in a small gap between two silver plugs

[1] O. J. Lodge, *Journal I.E.E.*, 19. p. 346. 1890.

in an exhausted and sealed glass tube (Fig. 369). A battery in series with the coherer produces a sufficient current, when the resistance of the coherer drops, to close a relay which actuates a Morse inker for recording the signals. A tapper of the form used with an electric bell gives a slight blow to the coherer to restore it to its original high resistance.

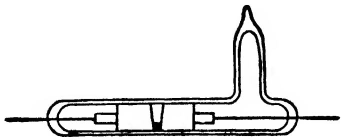


FIG. 369.

An arrangement of a receiving station is given in Fig. 370. By means of an inductance and sliding contact, P, the antenna circuit is tuned to resonance with the arriving electromagnetic waves. *a* and *b* are the primary and secondary coils of a transformer, and, the circuit C_1a being tuned to the antenna, oscillations are set up which induce oscillations in the coherer, the drop in whose resistance enables the cell B to produce sufficient current to attract the armature of the relay R, and actuate the Morse recorder M. Owing to the

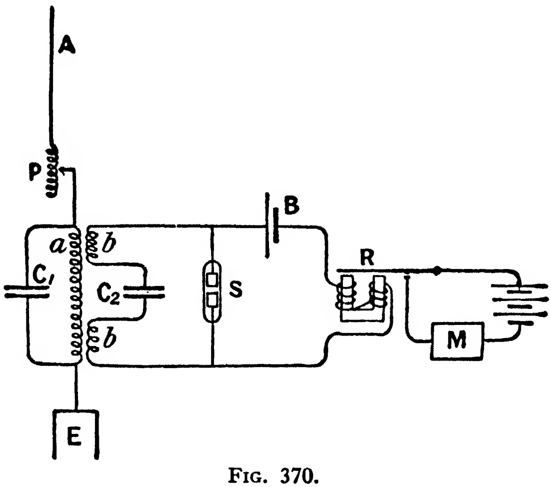


FIG. 370.

presence of the condenser C_2, which bisects the secondary coil of the transformer, there is no appreciable steady current in the battery and relay circuit, except when the coherer S has its low resistance due to the arrival of the electromagnetic waves.

Other methods have also been employed for the detection of electromagnetic radiations, amongst which we should note the employment of the heating effect in a fine wire, of the oscillatory current, the effect upon the hysteresis in iron and steel, and the oscillation valve. The triode valve displaced all the earlier forms of detector.

The Fleming Oscillation Valve (see p. 585) has also been used for detecting radiations. When the filament of the incandescent lamp (Fig. 371) is glowing, a current will only pass in the galvanometer circuit in such a direction that a negative charge flows from the filament to the plate P. Hence when the oscillatory current is brought in by the terminals A and B, it is unilateral in the galvanometer circuit, one-half of each oscillation being quenched by the lamp. A galvanometer deflection is then produced by the oscillations. The galvanometer may be replaced by a telephone, in which case each train of oscillations produces a sound in the telephone. Marconi has modified the arrangement by placing the telephone in series with the secondary circuit of a transformer, the primary of which is in the valve circuit. For the three-electrode valve see p. 585.

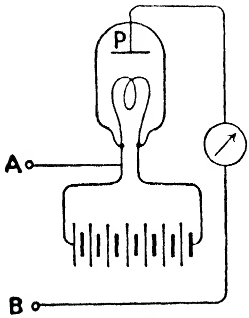


FIG. 371.

Crystal Detector.—A similar rectifying action is found in the case of certain crystals such as zincite, silicon, galena, carborundum, hertzite, etc. The crystal is fixed in a metal mount, and one of the sharp angles of the crystal bears against a metal plane, or a fine metal point bears upon the crystal. The conductivity of the contact depends upon the direction of the current.

Propagation of Waves in Wires and Cables.—The inductance and capacity of a single wire, although very small, become of importance when the currents in it vary with great rapidity, as, for example, when high frequency oscillatory currents are flowing in them, or when they are carrying fairly rapidly alternating currents, as in the case of those used in telephony.

Let L be the inductance per unit length of the cable, and R its resistance, then Ldx and Rdx are the inductance and resistance of

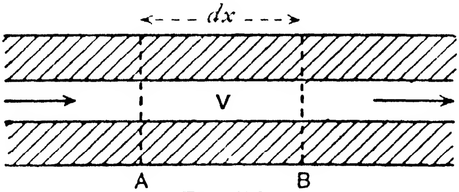


FIG. 372.

length dx. If V is the potential at the mid-point of the element, then $V-\frac{dV}{dx}\cdot\frac{dx}{2}$ is the potential at the end A of the section (Fig. 372) where the current enters, and $V+\frac{dV}{dx}\cdot\frac{dx}{2}$ at B, where

the current leaves. It may be noted that if x is measured from left to right, that is, in the direction of the current, then, if the cable has resistance only, $\frac{dV}{dx}$ is an essentially negative quantity. The p.d. between A and B is then $-\frac{dV}{dx}dx$. If the mean current in the element is I, the equation on p. 307 becomes

$$\mathrm{L}dx\frac{d\mathrm{I}}{dt}+\mathrm{R}dx\mathrm{I}=-\frac{d\mathrm{V}}{dx}dx$$

or,

$$\mathrm{L}\frac{d\mathrm{I}}{dt}+\mathrm{RI}=-\frac{d\mathrm{V}}{dx} \quad . \quad . \quad . \quad . \quad . \quad (\mathrm{i})$$

Again if the insulation of the cable leaks, the charge passing from the cable to the outer earthed sheath in time dt is $\mathrm{VK}dx\,dt$, where K is the conductance of the insulation from core to sheath for unit length of cable, and V is the mean positive potential of the section. K is sometimes called the *leakance*. If $\mathrm{I}-\frac{d\mathrm{I}}{dx}\cdot\frac{dx}{2}$ is the current entering the section at A, and $\mathrm{I}+\frac{d\mathrm{I}}{dx}\frac{dx}{2}$ that leaving it at B, the excess of charge leaving the section over that entering in time dt is $\frac{d\mathrm{I}}{dx}dx\,dt$. The two losses together will involve a drop in the potential of the section by the amount

$$\frac{\mathrm{VK}\,dx\,dt+\frac{d\mathrm{I}}{dx}dx\,dt}{\mathrm{C}\,dx},$$

where C is the capacity of unit length of the cable. In time dt this drop in potential may be represented by $-\frac{d\mathrm{V}}{dt}dt$. Thus,

$$\frac{\mathrm{VK}\,dx\,dt+\frac{d\mathrm{I}}{dx}dx\,dt}{\mathrm{C}dx}=-\frac{d\mathrm{V}}{dt}dt$$

$$\therefore \mathrm{C}\frac{d\mathrm{V}}{dt}+\mathrm{VK}=-\frac{d\mathrm{I}}{\cancel{dt}\,dx} \quad . \quad . \quad . \quad . \quad . \quad (\mathrm{ii})$$

Equations (i) and (ii) determine the current and potential and their variations at all points of the cable.

The case of most importance is that in which V and I vary harmonically, so that $\mathrm{V}=\mathrm{V}_0\cos pt$ is the real part of $\mathrm{M}\epsilon^{jpt}$, and $\mathrm{I}=\mathrm{I}_0\cos(pt+\theta)$ is the real part of $\mathrm{N}\epsilon^{jpt}$. The periodicity is the same for both, but the phases of V and I are not necessarily the same because M and N may be complex quantities (p. 378).

Equations (i) and (ii) may now be written

$$L\frac{d}{dt}(N\epsilon^{jpt})+RN\epsilon^{jpt}=-\frac{d}{dx}(M\epsilon^{jpt})$$

$$C\frac{d}{dt}(M\epsilon^{jpt})+KM\epsilon^{jpt}=-\frac{d}{dx}(N\epsilon^{jpt})$$

or,

$$\frac{d}{dx}(M\epsilon^{jpt})=-(jLp+R)N\epsilon^{jpt} \quad \ldots \ldots \quad \text{(iii)}$$

$$\frac{d}{dx}(N\epsilon^{jpt})=-(jCp+K)M\epsilon^{jpt} \quad \ldots \ldots \quad \text{(iv)}$$

Differentiating (iii) with respect to x and substituting from (iv), we get—

$$\frac{d^2}{dx^2}(M\epsilon^{jpt})=-(jLp+R)\frac{d}{dx}(N\epsilon^{jpt})$$

$$=(jLp+R)(jCp+K)M\epsilon^{jpt}$$

or,

$$\frac{d^2}{dx^2}(M\epsilon^{jpt})=P^2(M\epsilon^{jpt})$$

where P^2 is written for $(jLp+R)(jCp+K)$.

This equation determines the manner in which the quantity $M\epsilon^{jpt}$ varies with the distance of the point considered from the end of the cable.

A solution may be found as on p. 23 in the form $M\epsilon^{jpt}=\epsilon^{ax}$.

On differentiating twice and substituting

$$a^2\epsilon^{ax}=P^2\epsilon^{ax}$$

and

$$a=\pm P$$

The most general form of the solution is therefore

$$M\epsilon^{jpt}=A\epsilon^{Px}+B\epsilon^{-Px}$$

Similarly on differentiating (iv) and substituting from (iii)

$$N\epsilon^{jpt}=A_1\epsilon^{Px}+B_1\epsilon^{-Px}$$

where A, B, A_1 and B_1 are arbitrary constants. Suppose that the values of V and I at the origin of the cable are the real parts of $\overline{M}\epsilon^{jpt}$ and $\overline{N}\epsilon^{jpt}$. Then A and A_1 must both be zero, or the potential and current would become infinite at a great distance from the origin. But from the dissipation of charge and energy it follows that they must die away in travelling along the cable.

Thus, $\quad B=\overline{M}\epsilon^{jpt}, \quad$ and, $\quad B_1=\overline{N}\epsilon^{jpt}$

and, $\quad M\epsilon^{jpt}=\overline{M}\epsilon^{-Px}\epsilon^{jpt}, \quad N\epsilon^{jpt}=\overline{N}\epsilon^{-Px}\epsilon^{jpt},$

$$M=\overline{M}\epsilon^{-Px}, \qquad N=\overline{N}\epsilon^{-Px}.$$

The quantity P, or $\sqrt{(jLp+R)(jCp+K)}$, is itself complex, and may be written in the form $a+jb$.

Whence, $$M\epsilon^{jpt}=\overline{M}\epsilon^{-(a+jb)x}\epsilon^{jpt}$$
$$N\epsilon^{jpt}=\overline{N}\epsilon^{-(a+jb)x}\epsilon^{jpt}$$

or, $$M\epsilon^{jpt}=\overline{M}\epsilon^{-ax}\epsilon^{j(pt-bx)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (v)$$
$$N\epsilon^{jpt}=\overline{N}\epsilon^{-ax}\epsilon^{j(pt-bx)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (vi)$$

Now V and I are the real parts of $M\epsilon^{jpt}$ and $N\epsilon^{jpt}$. They are therefore given by the equations,

$$V=V_0 \cos pt=\overline{V}\epsilon^{-ax} \cos (pt-bx) \quad . \quad . \quad (vii)$$

and, $$I=I_0 \cos (pt+\phi)=\overline{I}\epsilon^{-ax} \cos (pt-bx+\phi) \quad (viii)$$

where $\overline{V}$ and $\overline{I}$ are the amplitudes at the origin and ϕ is their difference of phase, unknown at the moment.

We therefore see that their amplitudes V_0 and I_0 at any point are $\overline{V}\epsilon^{-ax}$ and $\overline{I}\epsilon^{-ax}$, and their phases are bx behind those at the origin. Hence, points for which x differs by the amount $\frac{2\pi}{b}$ are in the same phase at the same time, and $\frac{2\pi}{b}$ may be looked upon as the wave-length of the disturbance.

Further, we see that the amplitude $\overline{V}\epsilon^{-ax}$ diminishes exponentially as we pass away from the origin.

Again, the wave-length being $\frac{2\pi}{b}$, and the frequency n, $\frac{p}{2\pi}$ (p. 342),

$$\text{velocity}=\text{wave-length}\times\text{frequency}$$
$$=\frac{p}{b}.$$

The values of V are drawn in Fig. 373, for the instant at which

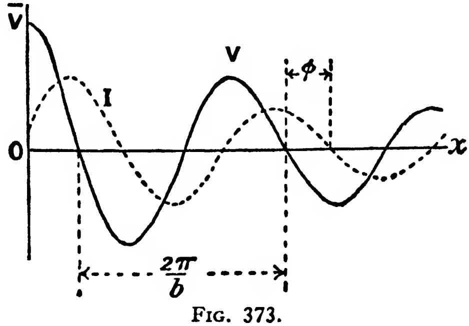


FIG. 373.

that at the origin is $\overline{V}$. As each point goes through a harmonic

change in the value of V, the points of zero value will now travel to the right with velocity $\frac{p}{b}$. This is the rate at which any given phase of the disturbance travels along the cable.

The relation between N and M may be obtained from equation

$$M\epsilon^{jpt}=\overline{M}\epsilon^{-Px}\epsilon^{jpt}.$$

for,

$$\frac{d}{dx}(M\epsilon^{jpt})=-P\overline{M}\epsilon^{-Px}\epsilon^{jpt}.$$

But from equation (iii)—

$$\frac{d}{dx}(M\epsilon^{jpt})=-(jLp+R)N\epsilon^{jpt},$$

$$\therefore\ N=\frac{P}{jLp+R}\overline{M}\epsilon^{-Px},$$

or,

$$N=\frac{P}{jLp+R}M,$$

and since,

$$P=\sqrt{(jLp+R)(jCp+K)},$$

$$N=\frac{\sqrt{jCp+K}}{\sqrt{jLp+R}}M.$$

Now,

$$jCp+K=\sqrt{C^2p^2+K^2}\epsilon^{j\theta_1}\qquad\text{(see p. 377)}$$

where,

$$\tan\theta_1=\frac{Cp}{K},$$

and,

$$jLp+R=\sqrt{L^2p^2+R^2}\epsilon^{j\theta_2},$$

where,

$$\tan\theta_2=\frac{Lp}{R},$$

$$\therefore\ \frac{jCp+K}{jLp+R}=\left(\frac{C^2p^2+K^2}{L^2p^2+R^2}\right)^{\frac{1}{2}}\epsilon^{j(\theta_1-\theta_2)},$$

and,

$$N=M\left(\frac{C^2p^2+K^2}{L^2p^2+R^2}\right)^{\frac{1}{4}}\epsilon^{j\frac{\theta_1-\theta_2}{2}}.$$

The current amplitude is therefore $V_0\left(\frac{C^2p^2+K^2}{L^2p^2+R^2}\right)^{\frac{1}{4}}$ and leads the potential, in phase, by the angle $\phi=\frac{\theta_1-\theta_2}{2}$. This is positive if the capacity effect predominates and negative if the inductance effect predominates. Thus, from equation (viii) the current leads the potential if ϕ is positive, and lags behind it if ϕ is negative.

The dotted curve in Fig. 373 indicates the current at every point of the cable. Its equation is—

$$I=\overline{V}\left(\frac{C^2p^2+K^2}{L^2p^2+R^2}\right)^{\frac{1}{4}}\epsilon^{-ax}\cos(pt-bx+\phi).$$

The quantity a is called the *attenuation factor*, since it determines the rate of decay of the amplitude of the oscillation as we pass along the cable; while b is called the *wave-length factor*, since $\frac{2\pi}{b}$ is the distance between points at which the phase at any instant is the same.

To determine a and b in terms of the constants of the conductor, we must remember that

$$\mathrm{P}=\sqrt{j\mathrm{L}p+\mathrm{R}}\sqrt{j\mathrm{C}p+\mathrm{K}}=a+jb.$$

Squaring and multiplying out, we have—

$$-\mathrm{LC}p^2+\mathrm{KR}+j(\mathrm{KL}p+\mathrm{RC}p)=a^2-b^2+2jab,$$

$$\therefore\ a^2-b^2=\mathrm{KR}-\mathrm{LC}p^2,$$

and, $$2ab=p(\mathrm{KL}+\mathrm{RC})$$ (see p. 377).

Substituting in the first, the value of b found from the second, we have—

$$a^2-\frac{p^2(\mathrm{KL}+\mathrm{RC})^2}{4a^2}=\mathrm{KR}-\mathrm{LC}p^2,$$

a quadratic in a^2, the roots of which are given by

$$2a^2=\pm\sqrt{(\mathrm{L}^2p^2+\mathrm{R}^2)(\mathrm{C}^2p^2+\mathrm{K}^2)}+(\mathrm{KR}-\mathrm{LC}p^2).$$

Since a^2 must be positive, a being a real quantity, the positive sign is taken. Similarly we may solve for b.

Then—

$$2a^2=\sqrt{(\mathrm{L}^2p^2+\mathrm{R}^2)(\mathrm{C}^2p^2+\mathrm{K}^2)}+(\mathrm{KR}-\mathrm{LC}p^2) \quad . \quad . \quad \text{(ix)}$$

$$2b^2=\sqrt{(\mathrm{L}^2p^2+\mathrm{R}^2)(\mathrm{C}^2p^2+\mathrm{K}^2)}-(\mathrm{KR}-\mathrm{LC}p^2) \quad . \quad . \quad \text{(x)}$$

The problem of the propagation of electrical currents in cables was first solved by Lord Kelvin,[1] but the inductance and leakance were there left out of account. The matter was rectified by Oliver Heaviside, who obtained the equations (i) and (ii).

When L and K are omitted,

$$2a^2=\mathrm{RC}p=2b^2,$$

$\therefore$ from (vii) and (viii), $$\mathrm{V}=\overline{\mathrm{V}}\epsilon^{-\sqrt{\frac{\mathrm{RC}p}{2}}x}\cos\left(pt-\sqrt{\frac{\mathrm{RC}p}{2}}x\right),$$

and, $$\mathrm{I}=\overline{\mathrm{I}}\epsilon^{-\sqrt{\frac{\mathrm{RC}p}{2}}x}\cos\left(pt-\sqrt{\frac{\mathrm{RC}p}{2}}x\right).$$

The velocity of propagation is in this case $\sqrt{\frac{2p}{\mathrm{RC}}}$, and the attenuation constant $\sqrt{\frac{\mathrm{RC}p}{2}}$.

[1] Sir W. Thomson, *Mathematical and Physical Papers*, vol. 2.

Telephone Circuits.—The frequencies of oscillation used in telephony are limited by the range of the human voice. Thus several hundred per second is the order of frequency of most importance. If, in transmitting a complex wave such as that produced by the human voice, the simple waves into which it may be resolved are not all transmitted with the same attenuation, the quality of the wave received at the end of the cable will differ from that transmitted. Since the attenuation depends upon the frequency, the distortion produced may have a more disturbing effect than the actual dying away with distance of the amplitude of the wave.

In the case of submarine and underground cables, where the conducting wire is surrounded by a conducting sheet, and separated from it by some insulating material, the capacity is relatively great and the inductance small.

Then, $$2a^2=R\sqrt{C^2p^2+K^2}+(KR-LCp^2).$$

In this case it is an advantage to increase K, the leakage from the cable, for this will make the attenuation constant less dependent upon the frequency $\frac{p}{2\pi}$. A reduction in R will in all cases reduce the attenuation factor. Hence the advantage of making the cables to have as low a resistance as possible. We may consider the attenuation to be the result of the loss in energy of the wave on account of the ohmic resistance of the conductor in which the current is flowing.

It has been suggested by O. Heaviside that a distortionless cable might be constructed by increasing to a suitable extent the amount of leakance.

Amongst other suggestions we find that of making $\frac{R}{L}=\frac{K}{C}$, in which case equation (ix) becomes

$$a^2=\frac{CR^2}{L}=\frac{LK^2}{C}=KR,$$

and (x) becomes $$2b^2=LC\,.\,2p^2,$$

and, $$b=p\sqrt{LC}.$$

In this case the attenuation factor is constant, and the velocity $\frac{p}{b}$ is $\frac{1}{\sqrt{LC}}$. The waves in this case are transmitted without distortion.

In cables, R and C are the most important terms, and it is therefore necessary to increase L and K. The latter is easy, for it is only necessary to diminish the insulation. To increase L,

the method of adding inductances at intervals by introducing a number of turns surrounding an iron core is commonly adopted, and E. Soleri [1] and M. Miniotti [2] have suggested the use of a strand of iron wire in the cable, the high permeability of which produces the desired increase in inductance.

High Frequency Circuits.—When the frequency of oscillation in a circuit is of the order occurring in the case of Hertz's oscillators, the quantities L^2p^2 and C^2p^2 are so great in comparison with R^2 and K^2, that the latter may be neglected. The equations to the waves of potential and current are then very much simplified, for equations (ix) and (x) become

$$2a^2 = KR$$

and,

$$2b^2 = 2LCp^2$$

since in the equation for b, KR may be neglected in comparison with $2LCp^2$. Then from (vii) and (viii),

$$V = \overline{V}\epsilon^{-\sqrt{\frac{KR}{2}}x} \cos\left(pt - \sqrt{LCp^2}x\right),$$

and,

$$I = \overline{I}\epsilon^{-\sqrt{\frac{KR}{2}}x} \cos\left(pt - \sqrt{LCp^2}x\right),$$

from which we see that the velocity of propagation is now $\frac{1}{\sqrt{LC}}$, that is, it is the inverse of the oscillation constant ($\sqrt{LC}$) of the cable.

Lecher's Wires.—The method of employing the steady oscillations set up in a wire to find the velocity of propagation was first used by Sir Oliver Lodge (p. 443), who succeeded in showing that

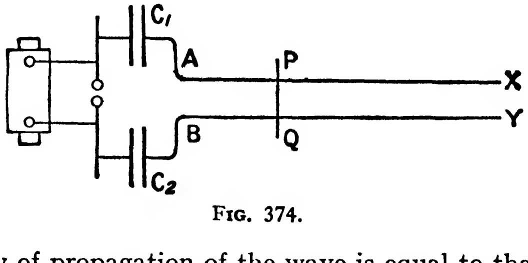


FIG. 374.

the velocity of propagation of the wave is equal to that of light. This method was also used by Hertz and afterwards modified by Sarasin and de la Rive, while E. Lecher [3] gave it the form shown in Fig. 374.

One coat of the condenser C_1 is connected to the wire AX, and one coat of C_2 to BY, the remaining coats being connected to the spark balls and to the induction coil. The variations of

[1] E. Soleri, *Atti dell' Assoc. Elettr. Ital.*, 12, p. 181. 1908.
[2] M. Miniotti, *Atti dell' Assoc. Elettr. Ital.*, 12, p. 193. 1908.
[3] E. Lecher, *Wied. Ann.*, 41, p. 850. 1890.

potential at C_1 start waves which travel down AX, and those at C_2 start similar waves in opposite phase down BY. It should be noted that in this case the wires are electrostatically coupled to the oscillator, in distinction to the magnetic or transformer coupling more frequently employed. There will consequently be a point of maximum variation of potential at each end of the wires, which therefore behave in an analogous manner to the open organ-pipe in the acoustical problem. Remembering that the wave along AX starts in opposite phase to that along BY, we may represent the fundamental steady vibration by the full-line curves in Fig. 375, and we see that nodes are situated at C and D. The first harmonic is given by the dotted curves, the nodes occurring at EF and GH. The nodes are found by placing a vacuum tube across the wires; this glows most brightly at the antinodes and ceases to glow at the nodes. A neon tube is most effective. If the conducting bridge PQ be placed across the wires, the variations in potential between X and Y will be a maximum when the bridge is situated at one of the nodes, in which case the wires may be looked upon as two circuits XPQY and C_1APQBC_2, having a common part PQ (Fig. 374).

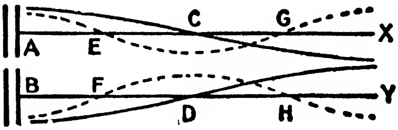


FIG. 375.

Wave-length or Frequency Meters.—There are now many forms of apparatus for measuring the frequency of oscillations or the wave-length of electric waves. The general principle is that a variable calibrated condenser is in series with a coil which serves as inductance and at the same time serves as a coupling with the circuit in which the oscillations occur. When the natural frequency of the meter is the same as that of the oscillations under investigation maximum resonance occurs. In order to detect the oscillations in the wavemeter a thermomilliammeter (p. 221) in its circuit may be employed, or a sensitive discharge tube of helium or neon may be connected across the condenser, as above.

Determination of Dielectric Constant by Oscillations.—The natural period of oscillation of a circuit, since it depends upon the capacity in the circuit, affords a means of determining the latter when the period of oscillation can be found. Sir J. J. Thomson[1] employed a parallel plate condenser, A (Fig. 376), near the plates of which are two flat conductors, E and F, to which the parallel wires EG and FH are connected. At each spark discharge at S, oscillations occur, and the periodic difference

[1] J. J. Thomson, *Phil. Mag.*, 30, p. 129. 1890.

of potential between E and F originates waves which travel down EG and FH. One of the spark knobs at S′ is connected to the wire EG at L, and the other to a movable contact M such that no spark occurs at S′. Then M is always in the same phase as L. Similarly, a neighbouring point N is found for the same condition

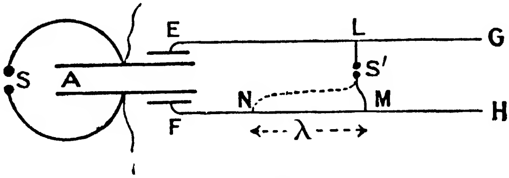


FIG. 376.

to be fulfilled. The distance MN is therefore one wave-length, since the phases at M and N are always the same.

The space between the plates A is now filled with the dielectric to be examined, and a new length M′N′ found, to correspond to the new wave-length λ'. Since the velocity of propagation of the wave in the wires is constant,

$$n\lambda = n'\lambda' = c.$$

But, $$n = \frac{1}{2\pi\sqrt{LC}}, \text{ and, } n' = \frac{1}{2\pi\sqrt{LC'}},$$

$$\therefore \frac{\lambda}{\lambda'} = \sqrt{\frac{C}{C'}}, \text{ and since } C_1 = kC, \frac{\lambda}{\lambda'} = \frac{1}{\sqrt{k}},$$

where k is the dielectric constant to be found. If the capacity C_1 of the rest of the circuit is also to be taken into account,

$$\frac{\lambda}{\lambda'} = \sqrt{\frac{C + C_1}{kC + C_1}}.$$

Also, using parallel wires, Lecher [1] found the length of the wires, which, together with the condenser, form a circuit which is in

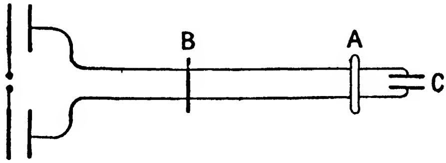


FIG. 377.

tune with a given source of oscillations. When the tube A containing a rarefied gas is laid across the Lecher wire it glows, unless the conducting bridge is laid upon the wires. It then ceases to glow, but for one particular distance, BC, the wires and

[1] E. Lecher, *Wied. Ann.*, 42, p. 142. 1891.

condenser form a circuit which " resounds " to the given oscillations and A again glows. The plates of the condenser C are adjustable, and on introducing the dielectric between them, A ceases to glow, but the plates are pushed together through a known distance until the initial value of the capacity is restored, as will be indicated by A again glowing. The dielectric constant of the medium may then be calculated as on p. 165.

The Heaviside and the Appleton Layers.—A problem in wireless telegraphy presented itself at an early date. It is the difficulty of seeing why electromagnetic waves should follow the curved surface of the earth. That the lower ends of the lines of force are anchored to the earth (p. 427) is a partial explanation. But it is not sufficient because there would still be attenuation in the upward direction. It was suggested by Heaviside that there is a conducting region in the atmosphere and that the waves are confined to the space between this region and the ground, very much as they are between Lecher wires (p. 459). Or the waves may be reflected one or more times from such a layer and so travel round the earth like the sound waves in a whispering galley. By observing reflections of impulses directed obliquely upwards, Appleton and others have shown clearly the existence of such a conducting region, the *ionosphere*, within which there are layers of maximum ionisation at different heights, of the order 80 to 300 km. The first layer discovered was named the *Heaviside layer*; there is also an *Appleton layer*. The heights of these layers and the intensities of the reflections have a marked diurnal variation, which points to their origin in electrons liberated by ultraviolet radiation (p. 473) from the sun. In corroboration of this, polarisation of the reflected waves can be explained in terms of the effect of the earth's magnetic field on the motion of the electrons.

Applications.—The uses of electromagnetic radiation in communication (wireless telegraphy and telephony) and in broadcasting (radio) are too widely known to require emphasis. The reflection of pulses of short-wave radiations from objects, such as aircraft, is the basis of radio-location or *radar*, invented in Great Britain originally for defence purposes. It is used also for navigational purposes and it is used to detect meteor streams, even in daylight, by means of the intensely ionised wake left by a meteor traversing the upper atmosphere.

CHAPTER XIV

CONDUCTION IN GASES

Spark Discharge.—The passage of an electric current across an air gap between two metallic conductors has been mentioned several times. At the atmospheric pressure, the difference of potential between the conductors required to start the current is quite different to that required to maintain it, and depends upon the shape of the electrodes employed. A spark point facilitates the discharge, as we should expect from the fact that the potential gradient in the neighbourhood is usually very great (p. 135). It is therefore necessary, when making measurements of sparking potential, to use for electrodes, polished spheres of diameter which is considerable in relation to the length of the spark-gap. The low resistance of the gap found in the experiments on oscillations bears no relation to the potential difference required to start the discharge, for when the current has passed for a short time, the gap is occupied by a quantity of highly conducting material derived partly from the gas and partly from the metallic electrodes.

It is therefore evident that Ohm's law is not applicable to the discharge; we must leave until later an examination of the relation between electromotive force and current.

The difference of potential between the electrodes required to start the discharge, and known as the spark potential, is independent of the metal of which the electrodes are made, except in the case of aluminium and magnesium, which metals have, under similar conditions, a less spark potential than the others, and for moderately great spark lengths the equation $V=a+bd$ represents fairly well the relation between spark potential V and spark length d. For the measurement of d, some form of spinterometer (Fig. 378) is employed, in which one of the knobs is made to travel by means of a micrometer screw. The potential difference is measured by means of an electrostatic voltmeter. The apparatus requires

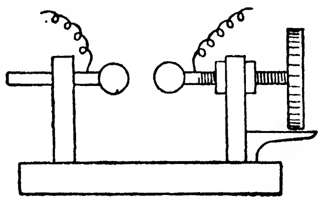


FIG. 378.

modification when the spark potential in a gas other than air is required.

As the spark becomes very short, the spark potential again increases, and has therefore a minimum, which occurs at some particular spark length whose value varies inversely as the gas pressure. At the atmospheric pressure this critical spark length is about 0·01 millimetre, and is therefore difficult to measure; but by lowering the pressure to about a millimetre of mercury, the critical spark length becomes of the order of 8 mm., and may then be easily measured. The curve in Fig. 379 indicates roughly the relation between V and d for very short gaps. The fact of the existence of the critical spark length may be shown by bringing the spark knobs together until their nearest points are at less than the critical distance apart. The spark will not then take place between the nearest points, but will move to a place where the distance apart of the spherical surfaces is the critical distance.

As the pressure is varied, the spark potential at first falls and then rises, a minimum occurring at some pressure called the

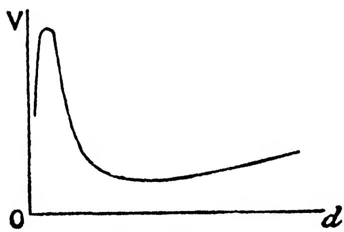


FIG. 379.

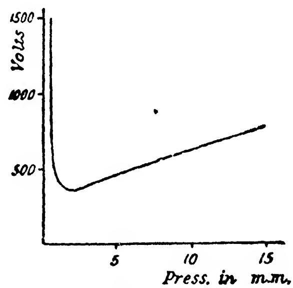


FIG. 380.

critical pressure. The mode of variation of potential difference and pressure for a moderate spark length of 3 mm. is shown in Fig. 380.

Paschen's Law.[1]—According to Paschen, the spark potential is proportional to the amount of gas between the knobs, that is, to the product of spark length and gas pressure. His measurements were all made at pressures above the critical pressure. Carr [2] has shown that the law holds good also at pressures below the critical pressure. Hence, if the relation between spark potential and pressure be known for one spark length, it may be calculated for all others.

Discharge at Low Pressure.—The measurements concerning the electric spark which are described above, do not involve

[1] F. Paschen, *Wied. Ann.*, **37**, p. 69. 1889.
[2] W. R. Carr, *Proc. Roy. Soc.*, **71**, p. 374. 1903.

any detailed knowledge of the processes going on, and we should probably still be in ignorance as to their nature if it were not for the Sprengel air-pump, the first to allow the attainment of very low pressures in glass bulbs. For very low pressures the mercury molecular pump is now used, and charcoal cooled by liquid air is used for absorbing most of the residual gases.

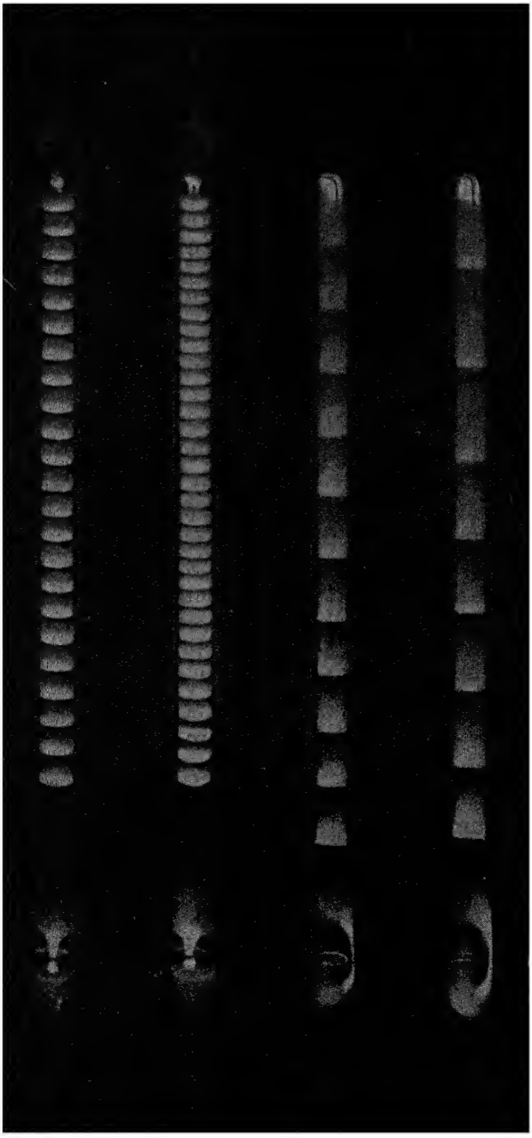
If the discharge between two conductors is maintained by means of a source of sufficiently high electromotive force, while the air pressure is continually reduced, the crackling nature of the discharge will after a time cease, and the path broadens out, giving a silent streamer whose colour varies with the gas employed, but generally changing from the white of the discharge at atmospheric pressure. At this stage a difference between the two ends of the discharge is noticeable, a discontinuity near the cathode being observable. The difference between the two ends becomes more accentuated as the pressure is further decreased, and the main part of the discharge will soon be seen to become stratified, and consist of layers of luminosity, separated by dark spaces. At a pressure of about 0·11 mm. of mercury, the discharge in hydrogen has the typical form shown in Fig. 381 (i), after De La Rue and Müller,[1] but the actual appearance of it cannot be described; it must be seen to be appreciated.

The discontinuity already observed near the cathode has increased considerably in size, and is called the Faraday dark space. Between it and the cathode is a luminous space called the negative glow, and between this and the cathode may now be seen a sharply defined dark space called the Crookes or cathode dark space. The positive column consisting of the striations extends from the Faraday dark space up to the anode.

On further reduction of the pressure, the scale of the phenomenon is enlarged, the growth taking place from the cathode, the positive column getting shorter and shorter and eventually disappearing as the Crookes dark space and the cathode glow expand. Fig. 381 (iv), shows the condition of the tube when there are still eight striations remaining, the pressure being then reduced to 0·037 mm. The phenomena at the cathode appear to be essential to the discharge, the positive column being accessory. At high pressures the separate parts of the discharge are so minute that their structure cannot be observed, but as the pressure is reduced, the mean free path of the molecules of the gas is larger, and the phenomenon of the discharge occurs on a larger and larger scale.

The boundary of the Crookes dark space is always luminous. When the boundary lies within the gas, we get there the cathode glow, but on reducing the pressure until the Crookes dark space

[1] W. De La Rue and H. W. Müller, *Phil. Trans.*, **169**, p. 155. 1878.



(i) (ii) (iii) (iv)

Fig. 381.

extends to the glass walls of the tube a bright phosphorescence is seen, the colour of which depends upon the nature of the glass of which the tube is made. It is a bright yellow-green in the case of soda glass, and a grey-blue for lead glass. Many minerals exhibit brilliant phosphorescence of various colours, when situated in this dark space.

The whole of the time that the pressure has been falling, the resistance of the tube has been decreasing. The fall of potential required to produce the discharge gets less and less, until the Crookes dark space reaches the walls of the tube ; but an increase in the discharge potential then begins, and at the highest vacuum attainable it is almost impossible to get a discharge through the tube.

Cathode Rays.—The phenomena occurring in the cathode dark space appear to be produced by something emitted by the cathode and travelling with great velocity, to which the name of cathode rays has been given ; they were investigated systematically by Sir William Crookes. We will here note, in addition to their power of exciting phosphorescence, some important properties.

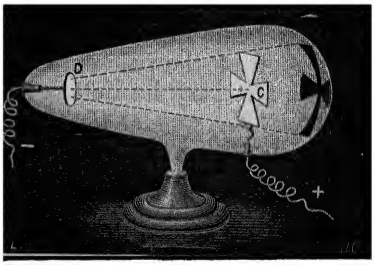


FIG. 382.

(i) The cathode rays travel in straight lines ; which fact may be observed by placing an obstacle in their path. Crookes placed a mica vane in the shape of a cross in the tube (Fig. 382), which will produce a dark shadow of its own shape upon the wall of the tube. By shaking down the cross after the phosphorescence has been produced for some time, the shape will still be seen, but it is now brighter than the surrounding parts of the glass, showing that after a time the glass surrounding the shadow has become " fatigued " by exhibiting the phosphorescence.

(ii) A body placed in the path of the rays experiences a mechanical force acting in a direction away from the cathode. If the rays impinge upon the upper face of the little mica mill wheel (Fig. 383), which is mounted upon an axle running upon horizontal rails, the wheel is rotated and may be driven from one end of the rails to the other. Sir J. J. Thomson has shown that the effect is not a purely mechanical one, the momentum of the rays being insufficient to produce the observed effect, but is probably due to the heating of the side of the mica upon which they impinge, the phenomenon being similar to that in the Crookes radiometer.

(iii) The rays produce heat when falling upon matter. If the cathode be concave in form, the rays being emitted normally from it are brought to a focus, and a thin piece of platinum or other substance may be raised to incandescence if situated at this point.

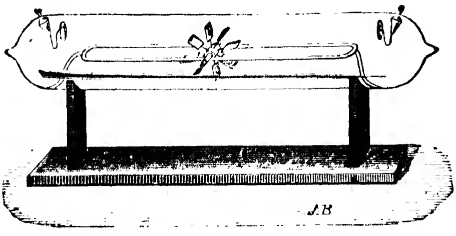


FIG. 383.

(iv) Cathode rays are deflected by a magnetic field exactly as an electric current would be; that is, they are moved in a direction at right angles to their own path and to the magnetic field. This effect may be exhibited by using a tube such as that shown in Fig. 386, where the beam of cathode rays may be deflected up or down by means of a magnetic field, the motion being observed by watching the position of the phosphorescent patch where the rays fall on the wall of the tube. On bringing the pole of a bar magnet near the tube, the beam becomes curved upwards or downwards according to the sign of the pole employed. The direction of deflection is such that it may be determined by the left-hand rule given on p. 240, *if the sign of the current from the cathode be taken as negative.*

(v) The rays are accompanied by a negative charge. Perrin [1] allowed the beam of cathode rays (Fig. 384) to pass into a hollow metallic cup, A, connected with an electrometer or electroscope. The instrument rapidly receives a negative charge; but a limit is soon reached owing to the gas in the discharge tube becoming conducting. If the sheath B be the cathode, and D the anode, A then acquires a positive charge, showing that positively charged bodies are moving in the opposite direction to the cathode rays. This point will be dealt with later.

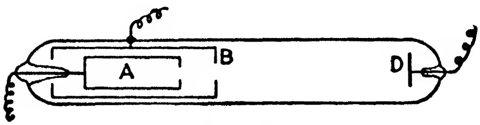


FIG. 384.

The above effects are explicable on the assumption that the cathode rays are streams of electrically charged particles, which acquire a very high velocity in the electric field maintaining the discharge. These bodies were called negative corpuscles by Sir J. J. Thomson, and subsequent investigation has shown that they are of very wide occurrence.

Wehnelt Cathode.—By substituting for the platinum electrode,

[1] J. Perrin, *Comptes Rendus*, **121**, p. **1130**. **1895**.

lime or one of the alkaline earths at high temperature, Wehnelt [1] showed that copious cathode rays can be obtained by means of comparatively low voltages. The lime or other oxide which is to form the cathode is placed upon a strip of platinum which is heated by means of an auxiliary current. Negative corpuscles are emitted by the oxide, which render it possible to send currents of 0·1 ampere through the tube with a p.d. of 100 volts. The corpuscles emitted have, on this account, comparatively low velocities of the order of 10^8 cm. per sec. (compare p. 472).

Röntgen or X-Rays.—It was observed by Röntgen that a fluorescent substance, situated near a discharge tube of high vacuum, exhibited luminescence as though exposed to ordinary light. Investigation showed that the emission which produced this effect proceeded from the walls of the vacuum tube upon which the cathode rays fell. Owing to their unknown nature Röntgen called them "X" rays. The name Röntgen rays has also been given to them.

For exhibiting luminescence, a sheet of cardboard covered with a layer of barium platino-cyanide forms a convenient arrangement. This renders the presence of the rays obvious. In addition to exciting luminescence in many substances, notably, the platino-cyanides, the Röntgen rays affect a photographic plate.

An important characteristic of the Röntgen rays is their penetrability for ordinary matter. The absorption of the rays by matter is dependent upon the density of the body upon which they fall, and hence the well-known application of the rays for producing shadows of the bones in the human frame, the flesh and portions of less density being the more transparent for them.

The rays produced in different vacuum tubes differ in character, the penetrating power being greater when the vacuum is higher. The highly penetrating rays from the tube of high vacuum are often spoken of as "hard" rays, and the relatively less penetrating rays from a tube of not so high a vacuum, as "soft" rays.

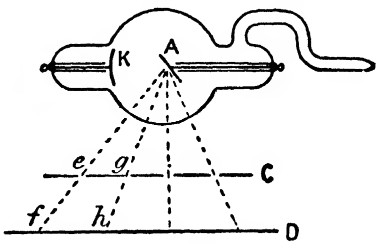


FIG. 385.

That the Röntgen rays arise at the point where the cathode rays strike a solid obstacle may be seen from the fact that the shadows cast by the rays arising where the cathode rays strike the walls of a vacuum tube are blurred; but if a concave cathode be used, and the cathode rays thereby focussed on a platinum plate A (Fig. 385), and a photograph obtained by placing a sheet of tinfoil, C, having a number

[1] A. Wehnelt, *Phil. Mag.*, 10, p. 80, July, 1905.

of pinholes made in it, over the photographic plate D, then, on developing D and replacing it, it will be found that the lines *fe*, *hg*, etc., when produced backwards, converge upon some point A, showing that the Röntgen rays proceeded from this point. The tube shown is of the original type employed for producing photographs, as the smallness of the source of the Röntgen rays renders it possible to produce shadows having extremely good definition.

Secondary Radiation.—When the Röntgen rays fall upon any material, other rays arise, to which the name of "Secondary Radiation" has been given. This secondary radiation is complex in character, and consists for the greater part of secondary X-rays. The secondary X-rays are of two distinct types, namely, scattered X-rays and characteristic X-rays. Scattered X-rays are of the same nature as the original beam, and may be considered as rays which have been deflected by the substance upon which they fall. Different substances do not scatter X-rays to the same extent, and there is no simple relation between the atomic weight of a substance and the amount of scattering it produces. With the lighter elements, the amount of scattering depends upon the mass present rather than the atomic weight, but the heavier elements scatter more than the lighter elements. The characteristic X-rays, as their name implies, vary in nature with the substance from which they arise. Their nature is of such great importance that it will be dealt with separately. Professor Barkla [1] found that all gases, upon which the Röntgen rays fall, emit secondary rays of the same penetrating power as the primary rays, and that the secondary rays from solid substances are sometimes polarised.

In addition to the scattered and the characteristic X-rays, it was found by Curie and Sagnac [2] that the secondary radiation consisted in part of negatively charged corpuscles moving with high velocity. It has been found that the intensity of emission of the corpuscular rays is greater the higher the atomic weight of the material. They are not emitted uniformly in all directions, being most freely emitted in a direction perpendicular to that of the X-ray beam. Their velocity does not appear to depend much upon the nature of the metal from which they arise, nor upon the distance of the X-ray tube from the material. For a number of different metals, including zinc, platinum and lead, the velocity of these negative corpuscles ranges between $6{\cdot}0 \times 10^9$ cm. per sec. and $8{\cdot}3 \times 10^9$ cm. per sec. It is thus about twice as great as those for the cathode rays given on p. 472.

[1] C. G. Barkla, *Phil. Mag.*, 5, p. 685 (1903); 11, p. 812 (1906); and *P. Roy. Soc.*, 77, p. 247 (1906).
[2] P. Curie and G. Sagnac, *Journ. de Physique*, 1, p. 13, Jan. 1902.

Innes[1] established the fact that the velocity of the corpuscular rays emitted by a body is independent of the distance of the body from the X-ray tube from which the primary rays arise. Further, the experiments of Beatty[2] and Sadler[3] show that the velocity of the corpuscular rays is equal to that of the cathode rays in the X-ray tube which gave rise to the primary X-rays.

Determination of Velocity and Ratio of Mass to Charge of the Corpuscles constituting the Cathode Rays.—On the assumption that the cathode rays consist of negatively charged corpuscles moving with high velocity, it becomes necessary to determine the three quantities, velocity, mass and charge associated with the corpuscle. The velocity and the ratio of mass to charge may be determined without great difficulty, but the determination of the actual mass and charge is more troublesome.

If e be the charge associated with the corpuscle, and v its velocity, we may consider it to constitute a current of strength ev. In a magnetic field of strength H, at right angles to the direction of motion, the force acting at right angles to both field and current is $\mathrm{H}ev$. A body which experiences a force always at right angles to its direction of motion describes a circular path, and the normal acceleration being $\frac{v^2}{r}$, where r is the radius of the path, the force is $\frac{mv^2}{r}$, m being the mass of the body. Hence the equation of motion for a corpuscle in a magnetic field is

$$\frac{mv^2}{r}=\mathrm{H}ev, \text{ or, } \frac{mv}{e}=\mathrm{H}r.$$

If, then, the stream of cathode rays produced by the cathode K (Fig. 386), and limited by the metal blocks A and B having

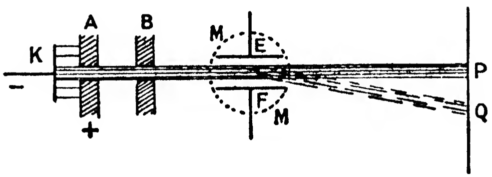


Fig. 386.

horizontal slots, pass through a magnetic field restricted to the circular space MM, then in passing through the field they will describe arcs of circles, the radius of which may be determined from the difference in position, PQ, of the luminous patch on the

[1] P. D. Innes, *Proc. Roy. Soc.*, A, 79, p. 442. 1907.
[2] R. T. Beatty, *Phil. Mag.*, 6, xx, p. 320. 1910.
[3] C. A. Sadler, *Phil. Mag.*, 6, xix, p. 337. 1910.

phosphorescent screen when the magnet field is on and when it is off. Thus H and r being known, the quantity $\frac{mv}{e}$ can be found.

Again, if the rays in their path have to traverse an electrostatic field due to the plates E and F maintained at a high difference of potential, the corpuscles experience a force eV while in the field, V being the electric intensity between the plates. If this electric field be at right angles to the magnetic field of the last experiment, and its intensity be arranged so that the force on the corpuscles is equal and opposite to that of the magnetic field, $eV=Hev$, or $\frac{V}{H}=v$; the corpuscle will now be undeviated so long as it is passing through the two fields. The fields are arranged to occupy the same part of the path, and are so varied in strength that the phosphorescent patch occupies its undisturbed position at the end of the tube. We then have $v=\frac{V}{H}$, and the velocity of the corpuscle is known. From the first experiment with the magnetic field alone, $\frac{mv}{e}$ is known, and therefore $\frac{m}{e}$ can be calculated.

By this method Sir J. J. Thomson,[1] to whom the method is due, obtained the following results :—

Gas.	v	$\frac{m}{e}$	Gas.	v	$\frac{m}{e}$
Air	$2\cdot8\times10^9$	$1\cdot3\times10^{-7}$	Air* . .	$2\cdot8\times10^9$	$1\cdot1\times10^{-7}$
Air	$2\cdot8\times10^9$	$1\cdot1\times10^{-7}$	Hydrogen .	$2\cdot5\times10^9$	$1\cdot5\times10^{-7}$
Air	$2\cdot3\times10^9$	$1\cdot2\times10^{-7}$	CO_2 . . .	$2\cdot2\times10^9$	$1\cdot5\times10^{-7}$
Air* . . .	$3\cdot6\times10^9$	$1\cdot3\times10^{-7}$			

* Platinum electrodes, the others being of aluminium.

The values of v vary, as would be expected, since v depends upon a number of conditions, but the values of $\frac{m}{e}$ do not differ very much from the mean, $1\cdot3\times10^{-7}$, which indicates that the corpuscles are of the same kind whatever the gas employed, or the metal used for electrodes.

This ratio $\frac{m}{e}$ given by this method, or the mass associated with unit charge, plays a part similar to that of the electro-chemical equivalent in electrolysis. Remembering that the same quantity for hydrogen is 0·0001045, we see that for the cathode rays the electro-chemical equivalent is of the order of $\frac{1}{1000}$ of that

[1] J. J. Thomson, *Phil. Mag.*, 44, p. 293. 1897.

for hydrogen. Three possibilities then present themselves: (i) there may be no simple relation between the mass of the hydrogen atom and of the corpuscle of the cathode rays on the one hand, or between the charges carried by them on the other; or (ii) if the masses are of the same order, the charge carried by the corpuscle is of the order of 1000 times that carried by the hydrogen ion in electrolysis; or (iii) if the charges are of the same order, the mass of the corpuscle is of the order of $\frac{1}{1000}$ that of the hydrogen atom. The question can only be settled by further experiment (see p. 482), but we may anticipate so far as to say that (iii) ultimately turned out to be near the truth. There is every reason to believe that the electric charge met with in the case of the electrolytic monovalent ion and in the corpuscle of the cathode rays is the ultimate and indivisible unit of electricity. Whether the corpuscle of the cathode rays is a small portion of "matter" with this charge associated with it, or whether it merely *is* the charge, is a question that we cannot enter into now. The name of *Electron* was suggested by Dr. Johnston Stoney for this fundamental unit of electrical charge first met with in the cathode rays, and the name is now universally adopted. We shall see presently that electrons are constituents of all matter, and play an important part in phenomena where their presence was unsuspected until after their discovery in the vacuum tube.

Method of Leakage in Ultra-violet Light.—It was found by Hallwachs [1] and others, that when ultra-violet light falls upon the negatively electrified surface of a sheet of zinc, the surface rapidly loses its negative charge; but if it be positively charged, there is no loss. This is explained if negative corpuscles are detached by the ultra-violet light from the surface, their repulsion from the negatively charged surface constituting the loss which is observed to take place. When the surface is positively charged the corpuscles are not driven away, and there is of course no loss.

Sir J. J. Thomson [2] made use of this phenomenon to determine the value of $\frac{m}{e}$ for these corpuscles, and found it to be of the same order of magnitude as for those in the cathode rays, which makes it presumable that the two are identical in kind.

A magnetic field, H, parallel to the negatively charged surface is maintained, and the paths of the corpuscles are thereby modified. Taking the value of the electric intensity due to the charge on the surface as V, the force on each corpuscle due to this field is Ve and is directed away from the surface, since the

[1] W. Hallwachs, *Wied. Ann.*, **33**, p. 301. 1888.
[2] J. J. Thomson, *Phil. Mag.*, **48**, p. 517. 1899.

charge of the corpuscle is negative. If the magnetic field H be directed from front to back (Fig. 387), the force on the corpuscle is $He v$ and is directed downwards. After leaving the surface the velocity will no longer be normal to it.

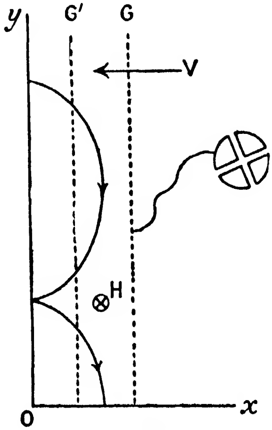


Fig. 387.

Taking the axis of x normal to the surface and the axis of y parallel to it and perpendicular to the magnetic field, the component of velocity parallel to Ox is $\frac{dx}{dt}$, and that parallel to Oy is $\frac{dy}{dt}$, and the corresponding accelerations are $\frac{d^2x}{dt^2}$ and $\frac{d^2y}{dt^2}$. If m be the mass of the corpuscles, the forces parallel to Ox due to the fields are Ve and H$e\,.\,\frac{dy}{dt}$, and these are together the resultant force $m\frac{d^2x}{dt^2}$ acting on the corpuscle in the x direction. Applying the left-hand rule of p. 240, and remembering that the moving corpuscle corresponds to a negative current, we see that the corpuscle when moving downwards experiences a force directed towards the plate, due to the magnetic field. The force equation for the components parallel to Ox is—

$$m\frac{d^2x}{dt^2}=\mathrm{V}e-\mathrm{H}e\frac{dy}{dt}.$$

Again, since there is no component of V parallel to Oy, we have for this direction the equation—

$$m\frac{d^2y}{dt^2}=\mathrm{H}e\frac{dx}{dt}.$$

The solution of these two simultaneous equations is—

$$y=\frac{\mathrm{V}}{\omega\mathrm{H}}(\omega t-\sin\omega t)$$

$$x=\frac{\mathrm{V}}{\omega\mathrm{H}}(1-\cos\omega t)$$

where $\omega=\frac{\mathrm{H}e}{m}$.

These are the equations of a cycloid formed by a circle rolling on the axis of y; for if P be the point on the circle when in the axis of y, and P′ the position of the point when the circle has

rolled through angle θ (Fig. 388), the length AP and the arc AP′ are equal, and the co-ordinates of P′ are therefore $x=a(1-\cos\theta)$ and $y=a\theta-a\sin\theta$. But if the circle roll with uniform angular velocity ω—

$$\theta=\omega t;$$

then,

$$x=a(1-\cos\omega t)$$
$$y=a(\omega t-\sin\omega t).$$

We see, then, that the moving corpuscle will describe a path similar to that of the point P upon the rolling circle, and its distance from the metallic surface will therefore never be greater than the diameter of the circle. By comparing the two sets of equations, we see that—

$$2a=\frac{2V}{\omega H}=2\cdot\frac{m}{e}\cdot\frac{V}{H^2}.$$

The conductor Oy is a zinc plate illuminated with ultra-violet light, for the liberation of the corpuscles. A parallel plate G (Fig. 387), connected to an electrometer, rapidly receives a negative charge when there is no transverse magnetic field, but with the field the corpuscles return to the plate Oy, and G does not receive any charge. In the position G′, the magnetic field does not affect the rate at which the charge is received. The limiting position is found, for which the field affects the rate at which G receives charge, and the distance between the plates is then $2a=2\frac{m}{e}\cdot\frac{V}{H^2}$. The limiting position is not so sharply defined as the equations indicate, but the mean value found for $\frac{m}{e}$ in this way is $1{\cdot}4\times10^{-7}$, which is in fair agreement with the result given on p. 472.

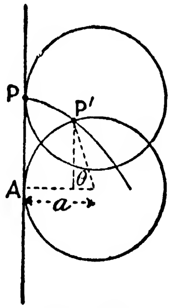


FIG. 388.

Ionisation.—Under ordinary circumstances, gases are very feeble conductors of electricity, a charged body situated in a gas retaining its charge for a very long time. Many agencies, however, render the gas a comparatively good conductor, amongst which may be mentioned, cathode rays, X-rays, hot bodies, flames and radio-active substances (Chap. XV). Further, the conductivity persists for a time, but does not last indefinitely. Sir J. J. Thomson and Lord Rutherford [1] showed that this conductivity may be removed in a variety of ways.

If the conductivity in the neighbourhood of the funnel A

[1] J. J. Thomson and E. Rutherford *Phil. Mag.*, **42**, p. 392. 1896.

(Fig. 389) be produced by means of an X-ray tube enclosed in a box covered with lead sheet to screen its direct effect from the electroscope, and provided with a window B, then the air drawn through the tube CD, into the electroscope by means of an aspirator will cause the leaves to collapse, whether the sign of the charge upon them be positive or negative.

A plug of glass wool placed in the tube at C will remove the conductivity from the air. The same result is produced if the conducting air be bubbled through water. Whatever it is that renders the air conducting is therefore filtered from it by these processes. The conductivity also disappears when an electric current passes through the air, which was shown by using for C a metallic tube having a wire stretched along its axis, a high potential difference being maintained between the tube and the wire. The leaves of the electroscope in this case do not collapse, showing that the cause of the conductivity has been removed.

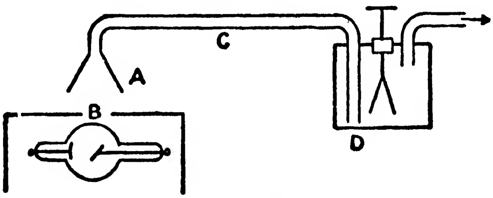


FIG. 389.

It is concluded from this experiment that the cause of the conductivity consists in charged particles, since they are driven to the sides of the tube or to the wire by the electric field, and further, that since the conducting gas as a whole does not exhibit electrification, the charged particles have opposite signs, and are in equal electrical quantities. These electrified particles are called *ions*, and the process of their production *ionisation*. It is now known that the ionisation of the gas is not the direct result of the X-rays, but of the corpuscular rays emitted by the gas (p. 483).

Conduction in Ionised Gas.—The conductivity of the ionised gas may be determined by bringing two parallel plates between which the gas is situated, to a known difference of potential, and the rate of change of potential of one of the plates determined by means of the quadrant electrometer.

If the capacity of the plate A (Fig. 390A) and the electrometer be known, the rise in potential per second enables the current passing from B to A to be determined. It is found, on gradually raising the applied difference of potential, that at first the current increases almost in accordance with Ohm's law, but the value of the current for further rise of potential difference falls below that

indicated by Ohm's law, and eventually a value is reached for which the current does not further increase. This current is known as the saturation current (Fig. 390B), and it is not exceeded until the electrical field is strong enough to produce ionisation in the gas. When this stage is reached the current begins to increase rapidly.

Saturation Current.—The saturation current depends upon the total number of ions between the plates, which in its turn depends upon the rate of production of ions and upon the volume of air between the plates. For the current i to pass from one plate to the other $\frac{i}{e}$ positive ions are driven against one plate, and $\frac{i}{e}$ negative ions against the other per second, and if q positive and q negative ions are produced by the Röntgen rays in one cubic

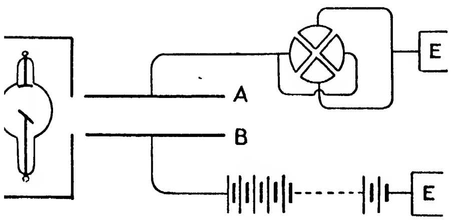


FIG. 390A.

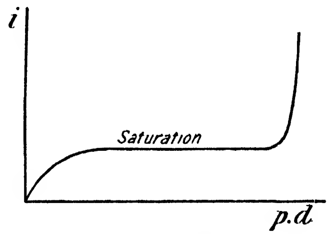


FIG. 390B.

centimetre per second, the total number of each kind produced per second in the space between the plates is qAl, where l is the distance apart of the plates and A the area of each, then $\frac{i}{e}$ cannot exceed qAl, and for the saturation current

$$qAl=\frac{i}{e}$$

or,

$$qAle=i.$$

The saturation current is therefore proportional to the distance apart of the plates, and we have the remarkable result that for the same difference of potential between the plates, the current increases if the plates are drawn farther apart. In their experiments on ionisation, Sir J. J. Thomson and Lord Rutherford[1] showed that when the ionisation is produced by Röntgen rays, this condition is realised.

Charge upon Negative Ion.—The problem of the determination of the charge of the negative ion produced in gases was successfully solved in 1898 by Sir J. J. Thomson. The current maintained in the gas by a known electrical intensity is measured. If U be the

[1] J. J. Thompson and E. Rutherford, *loc. cit.*

velocity of drift of the ions in electrical field of unit intensity, and E the actual electrical intensity, EU is the velocity of the ions. When the number present per c.c. is N and e the charge of each, NeEU is the current per sq. cm., which can be measured. To determine N, a discovery by C. T. R. Wilson[1] was used, which consists in the fact that in supersaturated, dust-free air, condensation takes place upon the ions, and a cloud of minute drops is formed. These drops fall through the gas at a constant rate, from which their size can be found, and knowing the total quantity of moisture condensed, from the work done in producing the adiabatic expansion necessary for the supercooling, N the number of drops can be found, and consequently e the charge associated with each ion is known.

Several methods have been employed for the determination of U, the velocity of drift of the ions produced by an electrical field of unit intensity, but that of Rutherford[2] is probably the most interesting. The ions are liberated from the zinc plate A (Fig. 391) by ultra-violet light from the source S, the plate being connected to a quadrant electrometer. The light passes through a window covered by a sheet of gauze B, and between A and B an alternating electromotive force is applied. During half a period of alternation the negative ions are driven away from A towards B, and during the next half-period they are driven back again. Whether they reach B or not depends upon its distance from A. If they do not reach B they will return to A, which will not lose a negative charge, and the electrometer deflection will not change, but if they do reach B they will not return, and A will continually lose negative charge. The distance between A and B is therefore adjusted until A begins to lose charge, and measured by the micrometer screw. This is the distance travelled by the negative ions during one half-period of the alternating electromotive force.

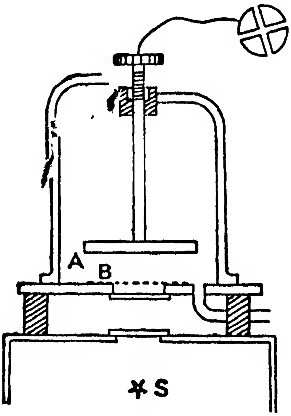


FIG. 391.

If d is the distance between A and B, and $e_0 \sin pt$ the alternating electromotive force between them, $\frac{e_0 \sin pt}{d}$ is the potential gradient, or electric intensity at any instant. U being the velocity of the ions for unit electric intensity, their instantaneous

[1] C. T. R. Wilson, *Phil. Trans.*, A, 189, p. 265. 1897.
[2] E. Rutherford, *Proc. Camb. Phil. Soc.*, 9, p. 401. 1898.

velocity is $\frac{Ue_0 \sin pt}{d}$. Taking x as the distance of any ion from A, $\frac{dx}{dt}$ is its velocity, and we have—

$$\frac{dx}{dt}=\frac{Ue_0 \sin pt}{d},$$

$$\therefore x=-\frac{1}{p}\cdot\frac{Ue_0}{d}\cos pt+C.$$

If $x=0$ when $t=0$, this means that the ion starts from the plate A,

and, $$C=\frac{Ue_0}{pd},$$

so that, $$x=\frac{Ue_0}{pd}(1-\cos pt).$$

Now $\cos pt$ varies between the values $+1$ and -1, and x is evidently a maximum when $\cos pt=-1$; and the greatest distance that the ion travels is—

$$\frac{2Ue_0}{pd}, \text{ or, } d=\frac{2Ue_0}{pd}, \qquad \therefore U=\frac{pd^2}{2e_0}.$$

Rutherford found the following values for U when $\frac{e_0}{d}$ is less than 1 volt per cm. For air, U=1·4 cm. per sec.; for hydrogen, 3·9 cm. per sec., and for CO_2, 0·78 cm. per sec.

Other methods both for the ions liberated from zinc by ultraviolet light and for those produced by Röntgen rays give practically the same result; which fact helps to establish the identity of the ions produced in these various ways.

The Condensation Experiments of C. T. R. Wilson [1] showed that if air saturated with water vapour and free from dust be suddenly cooled by causing an expansion exceeding 1 : 1·25 in volume, the vapour condenses upon the negative ions, but if the expansion exceeds 1 : 1·3 the condensation takes place upon the positive as well as the negative ions.

The explanation of the condensation upon the ions was given by Sir J. J. Thomson.[2] It is shown in text-books on the Properties of Matter that the maximum vapour pressure of water over a convex surface is greater than that over a plane surface by the amount δp, where $\delta p=\frac{2T\rho}{a(\sigma-\rho)}$, T being the surface tension of the liquid surface, a its radius of curvature, σ the density of

[1] C. T. R. Wilson, *Phil. Trans.*, A, **193**, p. 289. 1899.

[2] J. J. Thomson, "Applications of Dynamics to Physics and Chemistry," p. 165.

the liquid, and ρ that of the vapour. This change in the maximum vapour pressure is only significant if a is very small, but if such is the case the rise in maximum vapour pressure causes rapid evaporation of small droplets even in space that is saturated with respect to a plane surface. Formation and growth of a droplet is only possible when there is available a speck of dust or other suitable nucleus which presents a surface of sufficiently large radius of curvature to allow condensation with the degree of supersaturation present.

This change of vapour pressure with curvature can be attributed to the effect of the inward pressure $2T/a$ due to surface tension. Should a drop contain a charge e, there is an effective surface density of charge—

$$\sigma_1 = \frac{e}{4\pi a^2},$$

and the electrification will produce an outward pressure (p. 131) of—

$$\frac{2\pi\sigma_1^2}{k} = \frac{e^2}{8\pi k a^4}.$$

Subtracting this from the inward pressure due to surface tension gives the equation

$$\delta p = \frac{\rho}{\sigma - \rho}\left(\frac{2T}{a} - \frac{e^2}{8\pi k a^4}\right)$$

This gives the excess vapour pressure present, and hence the degree of supersaturation at a given temperature, in equilibrium with an ion of charge e and radius a. The presence of the charge is seen to reduce considerably the degree of supersaturation required to produce condensation for a given initial radius.

In the original apparatus used for these experiments, a quadrant electrometer was used as in Fig. 390A to measure the ionisation current flowing between the surface of some water in a bulb and the aluminium top cover of the bulb, in the presence of a Röntgen (X-ray) tube or other suitable ionising agent. Using the symbols of p. 478, this current, deduced from the rate of rise of potential dV/dt of the electrometer, of capacity C, is

$$C\frac{dV}{dt} = ANeEU,$$

where A is the horizontal cross-sectional area of the ionised column. From this can be found the produce Ne.

To find N, the number of ions per c.c. present, a cloud was induced to form on the ions by suddenly withdrawing air from the bulb through a side-tube, so that cooling occurred by adiabatic expansion. In the more modern apparatus of Fig. 392, the working chamber has as its floor the piston P, operated by a cam or similar device from a rotating handle, but in the original

arrangement the piston descended suddenly because the pressure below it was suddenly reduced. To prevent condensation occurring also on dust particles, the air was first cleared of dust by preliminary expansions, the dust being carried down by the droplets.

The cloud settled in a very regular manner, implying that the droplets were uniform in size, and the rate of fall of each droplet was deduced from observation of the rate of fall of the well-defined top of the cloud. It was shown by Sir George Stokes that the rate of fall of a spherical drop is $\frac{2}{9} \cdot \frac{ga^2}{\eta}$, where g is the acceleration due to gravity, a the radius of the drop, and η the coefficient of viscosity of the medium, here air, in which the drop falls. Thus the radius and therefore the volume of one drop is deduced.

The calculation of the lowest temperature attained in the adiabatic expansion requires a knowledge of the ratio in which the volume has increased, and this was deduced by applying Boyle's law, the pressure to which the gas finally rises after re-attaining the room temperature having been noted. After attaining this minimum temperature, condensation sets in. This ceases when the latent heat liberated has warmed the gas so far that the water still remaining in the vapour state just saturates the space. If K is the thermal capacity of 1 c.c. of the gas in these conditions, m_1 is the mass of water per c.c. in the supersaturated gas at the lowest temperature θ_1, and m_2 is the mass still as vapour at the subsequent temperature θ_2, m_1-m_2 must have condensed, evolving heat $L(m_1-m_2)$ where L is the latent heat of vaporisation at this temperature, so that $K(\theta_2-\theta_1)=L(m_1-m_2)$. m_1 is deduced from the original vapour content of the air, which was just saturated at the start, and m_2 must equal the saturation vapour density for temperature θ_2. Since m_2 is a function of θ_2, the solution of this equation, which is accomplished by trial, gives the required condensed mass m_1-m_2. As the mass of a single drop is now known, the number N of drops per c.c., and hence also of ions, is deduced. The electrical measurements had given Ne, whence e is known.

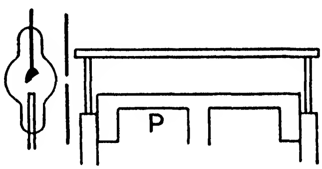


FIG. 392.

In his first experiments Sir J. J. Thomson obtained[1] the value $e=6{\cdot}5\times10^{-10}$ electrostatic units for ions in air, and $6{\cdot}7\times10^{-10}$ for ions formed in hydrogen. These expansions had not been enough to cause condensation on the positive ions as well as the negative (cf. p. 479), and in later experiments,[2] using radium as ionising source in place of the X-ray tube, which was in those

[1] J. J. Thomson, *Phil. Mag.*, **46**, p. 528. 1898.
[2] J. J. Thomson, *Phil. Mag.* (6th series), **5**, p. 346. 1903.

days somewhat unsteady in action, and using greater expansion ratios, he obtained $e = 3{\cdot}4 \times 10^{-10}$ e.s.u. or $1{\cdot}33 \times 10^{-20}$ e.m.e.

The Wilson Cloud Chamber.—From his studies of the condensation of droplets on ions, C. T. R. Wilson developed the Cloud Chamber to a stage where it yielded information of the utmost importance in Atomic Physics. Using strong illumination, instantaneous photographs were taken through the transparent top of the chamber, one example being reproduced in Fig. 393. This shows the appearance when the chamber just before or during the expansion is traversed by a narrow beam of X-rays. The droplets have formed on ions which are strung along the paths of swift particles which have been released from the gas molecules by the X-rays (p. 470); these particles are in fact electrons. The Wilson Cloud Chamber has been extensively

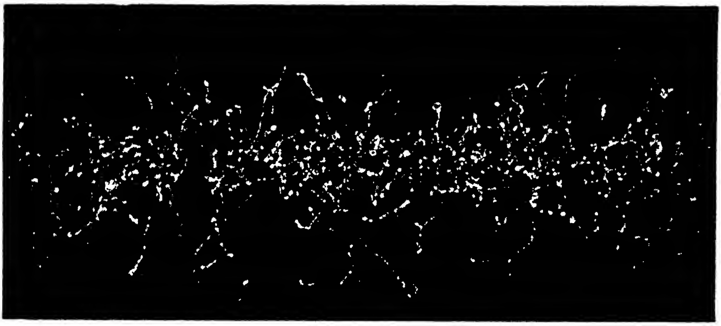


FIG. 393.

used in the study of the swift particles involved in radioactivity (Chapter XV) and in the investigation of cosmic rays (p. 509).

Charge of the Electron.—Professor Millikan modified the condensation method in such a way[1] that single charged drops could be observed and a number of the uncertainties of the cloud methods were removed. By using a liquid of low vapour pressure, usually an oil, evaporation from the drop during observation is prevented.

A fine spray of uniform droplets, produced by an atomiser, is blown into the air space above two parallel plates AB (Fig. 394). Five fine holes in the middle of the top plate admit drops to the space between the plates. An arc C illuminates the drops, which are observed by means of a long-focus microscope, an individual drop being singled out for observation. The drop is seen as a bright point of light, and the time it takes to fall under the action of gravity for a known distance is observed, from which its velocity v_1 is found.

[1] R. A. Millikan, *Phil. Mag.*, **34**, p. 1. 1917.

$$v_1=\frac{2}{9}\cdot\frac{ga^2}{\eta}(\sigma-\rho)$$

from Stokes' law (p. 481), where σ is the density of the drop and ρ the density of air. The effective weight of the drop is $w=\frac{4}{3}\pi a^3(\sigma-\rho)$, allowing for the buoyancy of the air.

$$v_1=\frac{2}{9}\cdot\frac{g(\sigma-\rho)}{\eta}\left(\frac{3w}{4\pi(\sigma-\rho)}\right)^{\frac{2}{3}},$$

or

$$w=\frac{4\pi(\sigma-\rho)}{3}\left(\frac{9\eta}{2g(\sigma-\rho)}\right)^{\frac{3}{2}}v_1^{\frac{3}{2}}.$$

X-rays are used to ionise the gas and battery D provides an electric field E between the plates. If the drop acquires a charge q it will experience a force Eq, up or down according to the sign of the charge. The resultant force is now $wg\pm Eq$. The resulting velocity v_2 is measured as before. The ratio of the velocities $v_1 : v_2$ is the ratio of the forces $wg : (wg\pm Eq)$, whence

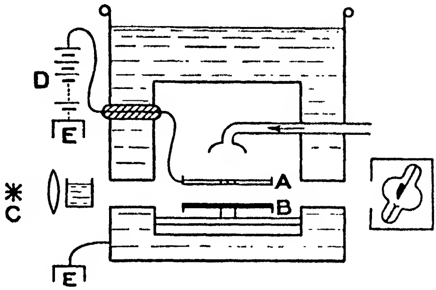


Fig. 394.

$$q=\pm\frac{v_2-v_1}{v_1}\cdot\frac{wg}{E}.$$

Substituting for w,

$$q=\pm\frac{4\pi}{3}\left(\frac{9\eta}{2}\right)^{\frac{3}{2}}\left(\frac{1}{g(\sigma-\rho)}\right)^{\frac{1}{2}}\frac{(v_2-v_1)v_1^{\frac{1}{2}}}{E}.$$

The velocity v_2 is found to change abruptly, corresponding evidently to the drop picking up or losing ionic charges, but the values found for q are all small integral multiples of one unit, which is therefore assumed to be the charge of a single electron.

Millikan found $e=4{\cdot}774\times10^{-10}$ electrostatic unit. Subsequent work, with a better value for the viscosity of air, gives

$$e=4{\cdot}802\times10^{-10}\text{ e.s.u.}=1{\cdot}602\times10^{-20}\text{ e.m.u.}$$

Combining this value of e with the value $1{\cdot}759\times10^7$ for e/m (in e.m.u.) gives $m=9{\cdot}11\times10^{-28}$ gm. e/m for the hydrogen ion in electrolysis is 9571 e.m.u./gm., and if the charge is the same, the ratio of the masses is

$$\frac{\text{H ion}}{\text{electron}}=\frac{1{\cdot}759\times10^7}{9571}=1835,$$

and the neutral hydrogen atom has a mass of $9{\cdot}11\times10^{-28}\times1836$ $=1{\cdot}674\times10^{-24}$ gm.

As the mass of 1 c.c. of hydrogen at 0° C. and 76 cm. pressure is $8{\cdot}987\times10^{-5}$ gm., this quantity contains $\dfrac{8{\cdot}987\times10^{-5}}{1{\cdot}674\times10^{-24}}=5{\cdot}38\times10^{19}$ atoms, or $2{\cdot}69\times10^{19}$ molecules. By Avogadro's hypothesis, this applies to any gas under the same conditions of temperature and pressure.

Electron Optics.—The paths of electrons in electric and magnetic fields may be likened to rays in optical systems. An arrangement of electrodes or coils which ensures that electrons passing through one point are then "focused" on another point, the "image" of the first, may be called an *electron lens* or *lens system*. Because the electron is so very small compared with the wavelength of light, a much higher resolving power is possible with an electronic system than with an optical one.

An electron of charge $-e$, starting with zero initial velocity from a cathode in a vacuum, will subsequently have a velocity v when at a position where the potential is V if $\frac{1}{2}mv^2=e\text{V}$, since its kinetic energy is acquired by virtue of the work done by the field. If it passes from a region of potential V_1 to one of V_2, the velocity changes from v_1 to v_2 where $v_2/v_1=\sqrt{(\text{V}_2/\text{V}_1)}$. This will involve in general a change of direction, *i.e.* a refraction. Assume that the change occurs abruptly, a case which may be closely approximated when an electron passes through an aperture in a wire-grid electrode. Let θ_1, θ_2 be the angles made respectively by the path of the electron with the normal to the surface separating the two regions. The only change of velocity will be at right angles to this surface, so resolving parallel to the surface,

$$v_1 \sin\theta_1=v_2\sin\theta_2$$

or,

$$\frac{\sin\theta_1}{\sin\theta_2}=\frac{v_2}{v_1}=\sqrt{\left(\frac{\text{V}_2}{\text{V}_1}\right)},$$

closely analogous to optical refraction, the refractive index being $\sqrt{(\text{V}_2/\text{V}_1)}$. In the more usual case, the effective refractive index varies continuously from place to place.

By use of suitably curved equipotential surfaces, usually coaxial cylinders and perforated discs of metal, a lens system may be simulated, as in Fig. 395A. Here the electrons emanate from a heated filament and the broken lines indicate typical paths, while some equipotential surfaces are shown. Relative to the filament or cathode, G is negative in potential and AB is positive. The combined effect of the repulsion of the electrons by G and their attraction by the nearer end of A is equivalent to the action of a converging lens. C is at a higher potential than AB; the effect here is that of a further converging lens followed by a weaker

diverging lens. The final image is formed on D, which may be a fluorescent screen or a photographic plate.

A magnetic field H, parallel to the axis of the system, will have effect only by virtue of any component velocity an electron may

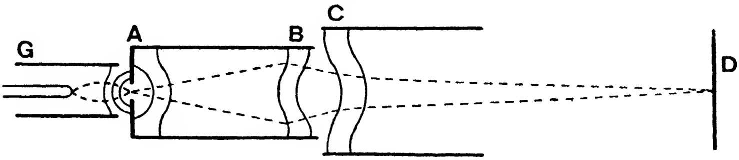


FIG. 395A.

have transverse to the field. For an electron leaving point A (Fig. 395B) with velocity v at an angle α to the field, there is a force $Hev \sin \alpha$ always at right angles to its path (p. 471). In the absence of the longitudinal component of its velocity, the electron would describe a circle of radius r where $mv \sin \alpha = Hev$. In fact it moves in a helical or screw-like path, completing one revolution round the cylinder on which the helix may be imagined to be wrapped in time $T = 2\pi r/v \sin \alpha = 2\pi m/eH$. In this time it moves a distance $l = Tv \cos \alpha = 2\pi mv \cos \alpha/eH$ parallel to the axis. Thus the field brings through D all electrons which pass through A and at A have the same component velocity parallel to the axis. This condition is attained if all the electrons have been accelerated by the same potential.

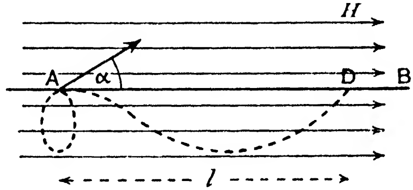


FIG. 395B.

As in the compound microscope, two or more images may be formed in succession. In the *electron microscope* the second image, or even the third, formed in this way, is received on a fluorescent screen or a photographic plate and very great magnification is possible and detail down to less than 10^{-7} cm. may be resolved.

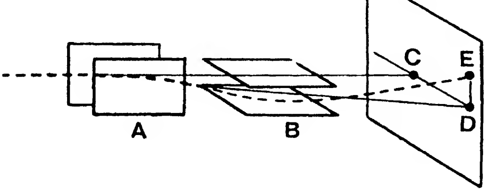


FIG. 395C.

Fig. 395C shows the principle of the electrode system used to deflect the focused spot on the screen of a *Braun* or *cathode-ray tube* from its central position at C. A potential difference between plates A moves the spot horizontally to D, while one across B produces a resultant deflection to E. Thus the spot traces out a graph connecting the two potential

differences, even when these vary rapidly. Many gauges and instruments are adapted to yield varying voltages as their output so that this technique may be used.

Absorption of X-rays.—It has already been seen (p. 469) that X-rays differ greatly in penetrating power, those from a "hard" tube being in general less absorbed by matter than those from a "soft" tube. This difference is associated with the greater velocity with which the cathode rays strike the anticathode, having been accelerated by the higher potential drop across a "hard" (high-vacuum) tube.

To discuss quantitatively the absorption, consider a beam of X-rays of intensity I to become $\mathrm{I}+d\mathrm{I}$ after traversing a thickness dx of a material: $d\mathrm{I}$ will of course be negative. The absorption may be written.

$$-d\mathrm{I}=A\mathrm{I}dx,$$

in which the quantity A, the *absorption coefficient*, is the fraction absorbed per unit length of path. If this coefficient is constant, we may integrate, and

$$\log \mathrm{I}=-Ax+\mathrm{C}.$$

Thus if $\mathrm{I}=\mathrm{I}_0$ for zero thickness, $\mathrm{C}=\log \mathrm{I}_0$ and the intensity after a thickness d is

$$\mathrm{I}=\mathrm{I}_0\epsilon^{-Ad}.$$

The intensity I is conveniently measured by the ionisation produced by the rays. To measure the absorption coefficient A, an ionisation chamber of the form shown in Fig. 396 may be used. A sheet of aluminium A hangs inside the chamber between parallel aluminium walls BB. A potential difference of several hundred volts is maintained between A and B, sufficient to produce the saturation current in the gas in the chamber. The rate of change of potential of A, which is a measure of the intensity I, is estimated by means of an electrometer or electroscope.

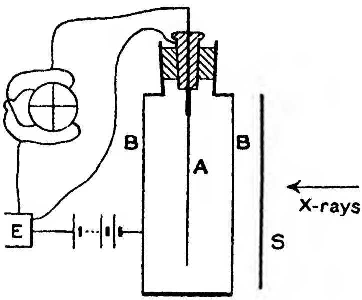


FIG. 396.

Measuring I_0, without screen S, and I with the screen, yields A. If the medium scatters the rays appreciably however, the scattering must be measured and allowed for.

The *mass absorption coefficient*, the ratio of the absorption coefficient to the density ρ, is in many ways a more useful quantity to use than the absorption coefficient itself, because it varies less on passing from one substance to another.

The absorption coefficients are only definite if the X-rays used are homogeneous, all of the same quality, which will be in-

terpreted later as implying that they must all have the same wave-length.

Characteristic X-rays.—By a study of the absorption of X-rays excited in different ways, it was found by Barkla and Sadler[1] that most substances emit one or more kinds of homogeneous secondary X-rays, characteristic of the substance and not of the quality of the primary rays, but that these rays cannot be excited by primary rays softer than themselves. They can in turn produce characteristic rays, of lower penetrability, from other substances.

Barkla,[2] using absorbing screens of aluminium, deduced from measurements of mass absorption coefficients for many substances

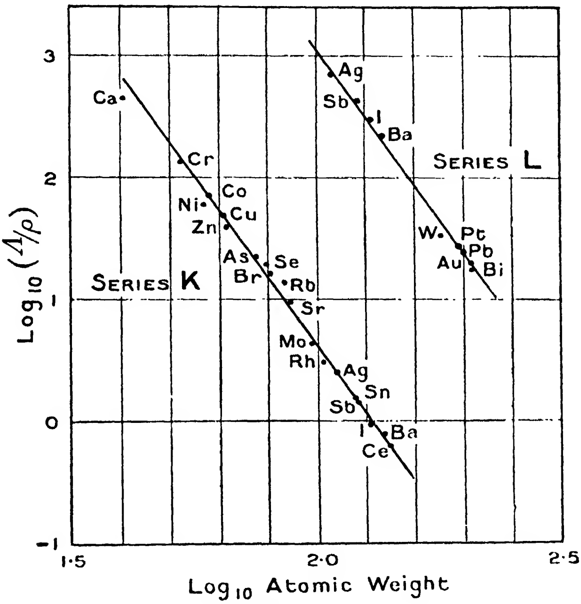


FIG. 397.

that the characteristic X-rays could be grouped into two distinct series, which he called the series K and series L fluorescent radiations. In some cases (Ag, Sb, I and Ba) a single substance emitted rays belonging to both series and in such cases the K radiation has a penetration (measured by the reciprocal of the absorption coefficient) some 300 times that of the L radiation. The mass absorption coefficient Λ/ρ, plotted against atomic weight of the absorbing element, gives two curves, but the data are better represented as in Fig. 397 where the logarithms of both quantities are plotted, for then two straight lines are obtained.

The lighter elements give very soft L radiations, absorbed

[1] C. G. Barkla and C. A. Sadler, *Phil. Mag.*, **16**, p. 550, Oct. 1908.
[2] C. G. Barkla, *Phil. Mag.*, **6**, xxii, p. 396. 1911.

heavily even in air, and the K series of the heavier elements could not be excited at the time of these experiments, there not being sufficiently hard primary radiation available, but it was later established that the characteristic radiations from the atoms of all elements can be classified in this same general way, into K, L, M, . . . series.

Direct bombardment of a substance by cathode rays excites, in general, homogeneous characteristic rays mixed with other rays, but Kaye[1] showed that subsequent passage of the radiation through a " filter " of the same material caused the absorption of most of the other radiation or its conversion to characteristic radiation so that an intense nearly homogeneous beam could be obtained.

Whiddington showed[2] that the minimum velocity V of the cathode rays required to excite the characteristic radiation of an element is given in cm. per sec. for the K and L radiations by the formulæ $V_K = 2(Z-2) . 10^8$ and $V_L = (Z-15) . 10^8$, where Z is the " atomic number," or ordinal number of the elements when they are arranged in a sequence, nearly the order of ascending atomic weights, to be discussed later (p. 496).

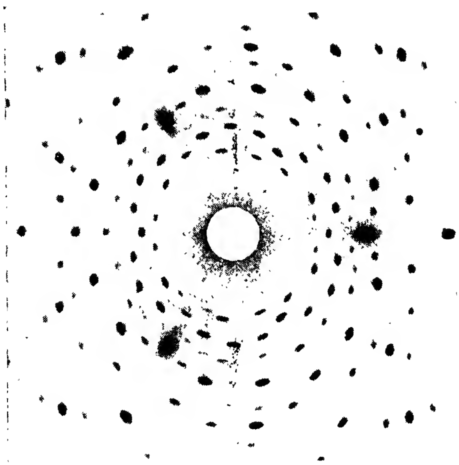
Interference and Reflection of X-rays.—X-rays are now known to be electromagnetic waves of very short wave-length, but a decisive proof of this was difficult to find until Prof. M. Laue suggested[3] that the atomic structure of crystals might form a sufficiently fine diffraction " grating " for this purpose. Friedrich and Knipping[4] interposed in turn various crystals (copper sulphate, rock salt, diamond, zinc blende) in a fine beam of X-rays falling on a photographic plate. Around the central spot there was in each case a pattern of spots, confirming Laue's theory. Fig. 398 shows the " Laue diagram " for quartz, the central heavy spot having been deliberately avoided.

The mathematical treatment by Laue of the production of the spots is very complex, but a much simpler procedure was soon introduced by W. L. (now Sir Lawrence) Bragg. Consider first the wave-front AB (Fig. 399A) meeting a plane CD containing regularly spaced particles. We assume, as in the Huyghens construction of ordinary optics, that some radiation is scattered from each particle, spreading out as a spherical wavelet. When the incident wave-front reaches D, the wavelet from C will have reached E and the wavelets from intermediate particles will at this instant all touch the plane DE, which is thus the reflected wave-front.

The penetration of X-rays is such that many such layers will be involved, and in general the waves from these will be out of phase

[1] G. W. C. Kaye, *Phil. Trans. Roy. Soc.*, A, **203**, p. 123. 1908.
[2] R. Whiddington, *Phil. Mag.*, **39**, p. 694. 1920.
[3] M. Laue, *Phys. Zeit.*, **14**, p. 421. 1913.
[4] W. Friedrich, P. Knipping, and M. Laue, *Le Radium*, **10**, p. 47. 1913.

and will interfere. In Fig. 399B is shown an advancing wave-front AB which meets parallel planes rich in atoms at B, C and E, and all the reflected waves travel in the direction EFG. If the path-difference between waves reflected from successive planes is



[By courtesy of the Research Laboratories, G.E.C.

FIG. 398.[1]

an integral number of wave-lengths, these waves will reinforce and intense reflection will occur, but in all other cases the set of successive waves will give more or less complete cancellation.

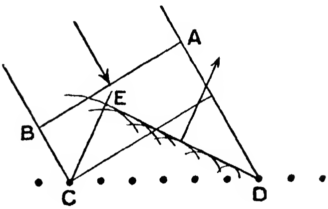


FIG. 399A.

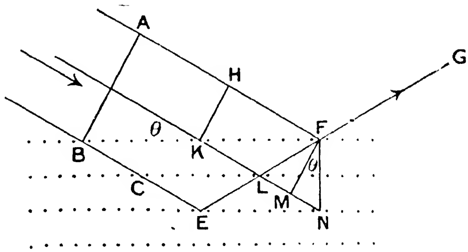


FIG. 399B.

The condition for reinforcement is

$$KL+LF-HF=n\lambda,$$

where n is integral. Produce KL to N making LN=KL=LF (which makes KFN a right angle) and draw FM perpendicular to LN.

Then $$KL+LF=KN$$

and, $$KM=HF.$$

[1] X-rays from molybdenum target, incident on quartz crystal parallel to the trigonal axis.

Hence

$$\begin{aligned}\text{difference of path} &= MN \\ &= FN \sin\theta \\ &= 2d\sin\theta\end{aligned}$$

where d is the perpendicular distance between the planes of scattering particles. Thus intense reflection occurs if

$$2d \sin\theta = n\lambda,$$

a result known as *Bragg's equation.* Note that θ here is the " glancing angle " between the incident beam and the *surface,* not the angle made with the normal.

On allowing a homogeneous beam of X-rays to fall at a glancing angle upon a face of a crystal, there is in general no reflected beam, but on rotating the crystal slowly, it is found that for certain angles θ_1, θ_2, θ_3, . . . a reflected beam may be detected. Then,

$$\frac{\lambda}{d} = 2\sin\theta_1 = \sin\theta_2 = \frac{2}{3}\sin\theta_3 = \text{etc.},$$

from which either λ or d may be found if the other is known. Other values for the ratio λ/d may be obtained by turning the crystal so that new sets of planes are employed. These planes are distinguished solely by the property of being rich in scattering particles (atoms, ions, etc.) and in general many such exist in a crystal. These planes need not be parallel to any particular face of the crystal and indeed the latter may even be an irregular fragment.

Refraction of X-rays.—The derivation given above of Bragg's equation assumes that the wave-length is unaltered within the crystal. It was found by Stenström[1] that the wave-lengths calculated from the observed deviations in the spectra of different orders, *i.e.* different values of n, did not agree exactly and he deduced that the rays are bent away from the normal when entering a crystal. Thus the refractive index is slightly less than unity, and may be written $\mu = 1 - \delta$.

Total internal reflection should therefore occur at a sufficiently large angle of incidence i (or small glancing angle θ, $\theta + i = \pi/2$), the critical glancing angle being given, as in optics, by $\sin i = \mu = \cos\theta$, and this in turn $= 1 - \frac{1}{2}\theta^2$ as θ is very small. Hence $\theta^2 = 2(1-\mu) = 2\delta$, $\theta = \sqrt{2\delta}$. Compton[2] found for crown glass a critical glancing angle of 11 minutes, giving $\delta = 5 \times 10^{-6}$.

Correction of Bragg's Equation.—Owing to refraction of X-rays on entering a crystal, it is the angle θ' (Fig. 400) inside the crystal which obeys the Bragg equation, but the external angle θ is the

[1] W. Stenström, *Dissertation*, Lund. 1919.
[2] A. H. Compton, *Phil. Mag.*, **45**, p. 1121. 1923.

one measured. The law of refraction in optics, translated into glancing angles, gives

$$\mu=\frac{v}{v'}=\frac{\lambda}{\lambda'}=\frac{\cos\theta}{\cos\theta'}=1-\delta,$$

where v and λ are the velocity and wave-length in air, v' and λ' in the crystal. Then $n\lambda'=2d\sin\theta'$, which may be written

$$n\frac{\lambda}{1-\delta}=2d\sqrt{\left\{1-\frac{\cos^2\theta}{(1-\delta)^2}\right\}}$$

$$n\lambda=2d\{(1-\delta)^2-(1-\sin^2\theta)\}^{\frac{1}{2}}$$

$$=2d\{\sin^2\theta-2\delta+\delta^2\}^{\frac{1}{2}}$$

$$=2d\sin\theta\left\{1-\frac{2\delta}{\sin^2\theta}\right\}^{\frac{1}{2}}$$

approximately

$$=2d\sin\theta\left(1-\frac{\delta}{\sin^2\theta}\right)\quad.\quad\text{(i)}$$

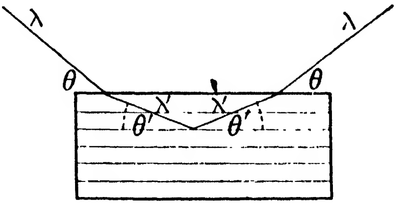


FIG. 400.

assuming δ to be very small.

Let measurements be made with one fixed wave-length λ but in two orders n_1, n_2 and let the apparent wave-lengths derived from the uncorrected Bragg equation be λ_1, λ_2, deduced from the glancing angles θ_1, θ_2 respectively. Then

$$n_1\lambda_1=2d\sin\theta_1\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad\text{(ii)}$$

while (i) becomes $\quad n_1\lambda=2d\sin\theta_1\left(1-\dfrac{\delta}{\sin^2\theta_1}\right)$

so that $\quad \lambda_1=\lambda\left(1-\dfrac{\delta}{\sin^2\theta_1}\right)^{-1}$

$$=\lambda\left(1+\frac{\delta}{\sin^2\theta_1}\right)\text{ approximately.}$$

Likewise $\quad \lambda_2=\lambda\left(1+\dfrac{\delta}{\sin^2\theta_2}\right)$

so that $\quad \lambda_1-\lambda_2=\lambda\delta\left(\dfrac{1}{\sin^2\theta_1}-\dfrac{1}{\sin^2\theta_2}\right).$

As λ_1 and λ_2 differ very little from λ, we may write

$$\frac{1}{\sin^2\theta_1}=\left(\frac{2d}{\lambda}\right)^2\cdot\frac{1}{n_1^2}$$

and $\quad \left(\dfrac{1}{n_1^2}-\dfrac{1}{n_2^2}\right)\delta=\dfrac{\lambda_1-\lambda_2}{\lambda}\left(\dfrac{\lambda}{2d}\right)^2.$

Thus $\quad \delta=\dfrac{\lambda_1-\lambda_2}{\lambda}\cdot\dfrac{n_1^2n_2^2}{n_2^2-n_1^2}\cdot\dfrac{\sin^2\theta_1}{n_1^2}$

$$=\frac{\lambda_1-\lambda_2}{\lambda}\cdot\frac{n_2^2}{n_2^2-n_1^2}\cdot\sin^2\theta_1,$$

a result known as Stenström's equation.

Once δ is known for a material, the true glancing angle θ' can be calculated for each observed angle θ and the wave-lengths in the crystal and in air can then be evaluated, assuming d to be known.

X-ray Spectrometer.—The intensity of the beam reflected in various directions from a crystal may be measured by an X-ray spectrometer. Fig. 401 illustrates the original form due to Sir William (Prof. W. H.) Bragg. Slits B, B′ in metal screens limit the beam coming from the target A. The position of the crystal C, mounted in wax on a table rotating about a vertical axis, is shown by vernier V. The arm D, rotating about the same vertical axis, carries the ionisation chamber E and a second vernier (not shown). The adjustable slit F can be used to limit the width of the reflected beam, which enters the chamber through a thin aluminium window. The chamber is insulated and raised to a high potential by a battery of cells, and the electrode inside it is connected to the gold leaf of a tilted electroscope (see Chapter XV), so that the motion of the leaf is a measure of the ionisation in the chamber and hence of the intensity of the reflected X-ray beam. To increase absorption of the X-rays, the chamber is usually filled with a suitable dense gas: sulphur dioxide gives about ten times the absorption of air, while with highly penetrating rays, methyl bromide, which absorbs still more, may be used.

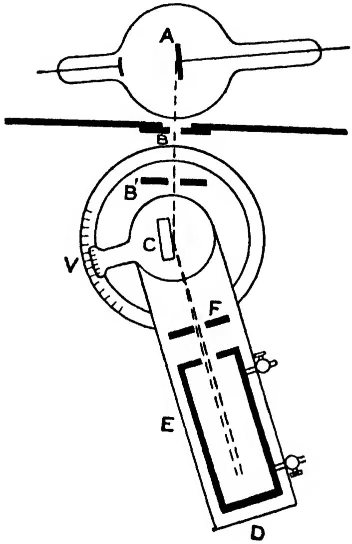


Fig. 401.

In plotting a spectrum starting with zero glancing angle, the crystal is turned at half the rate of the chamber and at each new setting the leaf of the electroscope is initially earthed and its motion then observed for a few seconds to estimate the intensity of the reflected rays. A typical plot of ionisation against angle is indicated in Fig. 402. The maxima A_1 and B_1 correspond to two characteristic radiations in the incident beam, which arise from a rhodium anti-cathode, the reflecting crystal being sodium chloride. The maxima A_2B_2 and A_3B_3 are produced by second-

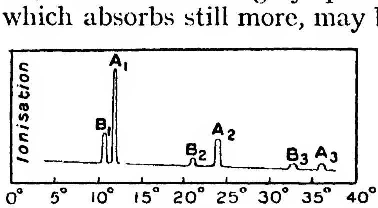


Fig. 402.

and third-order reflections. Such a curve is called an X-ray spectrum. The maxima A_1, A_2 and A_3 occur at angles which are nearly 11·8°, 23·5° and 36°,

$$\sin 11{\cdot}8^\circ : \sin 23{\cdot}5^\circ : \sin 36^\circ = 0{\cdot}204 : 0{\cdot}40 : 0{\cdot}63 = 1 : 2 : 3 \text{ nearly.}$$

Wave-length of X-rays and Crystal Structure.—The halogen salts, which crystallise on the cubic system, provide a simple illustrative example of the Bragg method. The investigation of these salts was of great importance, because it established the dimensions of certain crystalline structures and enabled the wave-lengths of certain characteristic X-rays to be determined in absolute measure. Using these distances and wave-lengths as reference standards, others could then be determined by comparative methods.

The simplest of arrangements conceivable for a crystal of the cubic system is that in which one atom is situated at each corner of a cube (Fig. 403A, p. 494). This structure repeated in all directions is called a *space-lattice*, and in this simple case there are three systems of planes rich in atoms, which may be easily recognised.

Let d_1 be the separation of planes such as ABCD and EFGH: this distance is the length of one edge of the basic cube. Planes parallel to AEGC are separated from one another by distances $d_2 = d_1/\sqrt{2}$. A third set of planes is parallel to ACF, which has a perpendicular distance from B (which lies in the next plane of the set) of d_3. Triangle ABK is drawn in true shape at (ii) in Fig. 403A, and it will be seen that

$$\frac{d_3}{d_1} = \frac{d_2}{\sqrt{(d_1^2 + d_2^2)}}$$

$$= \frac{d_1}{\sqrt{2}} \frac{1}{\sqrt{(d_1^2 + \frac{1}{2}d_1^2)}}$$

$$= \frac{d_1}{\sqrt{2}} \frac{\sqrt{2}}{d_1\sqrt{3}},$$

$$d_1 = d_3\sqrt{3}.$$

$$\therefore \frac{1}{d_1} : \frac{1}{d_2} : \frac{1}{d_3} = 1 : \sqrt{2} : \sqrt{3}.$$

Bragg obtained 5·22°, 7·30° and 9·05° for the respective reflections from a given crystal of potassium chloride. Using $\lambda = 2d \sin \theta$,

$$\frac{1}{d_1} : \frac{1}{d_2} : \frac{1}{d_3} = \sin 5{\cdot}22^\circ : \sin 7{\cdot}30^\circ : \sin 9{\cdot}05^\circ$$

so

$$\frac{1}{d_1} : \frac{1}{d_2} : \frac{1}{d_3} = 0{\cdot}0910 : 0{\cdot}1272 : 0{\cdot}1570 = 1 : \sqrt{2} : \sqrt{3}$$

in agreement with the assumption of a simple cubic lattice.

It should be noted that the atomic weights of potassium (39) and of chlorine (35·5) are not very different, so that the scattering powers of the two kinds of atoms may be expected to be similar. Thus in the above accounts no distinction has been made between them.

Sodium chloride gives different results, which may be interpreted by assuming that the sodium (Na=23) and chlorine atoms alternate in the lattice, as in Fig. 403B, where the sodium atoms are represented by black dots and the chlorine by white. Each set can be regarded as forming a lattice with an atom at each corner of a cube and one in the centre of each face. Writing $d_1=\frac{1}{2}EA$, $d_2=d_1/\sqrt{2}$; and both the corresponding planes contain atoms of both kinds. On the other hand, planes parallel to ACF contain atoms all of one kind, and the separation between planes of the same kind is $2d_1/\sqrt{3}$ instead of $d_1/\sqrt{3}$. The first-order

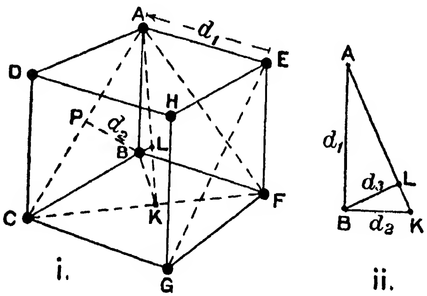


FIG. 403A.

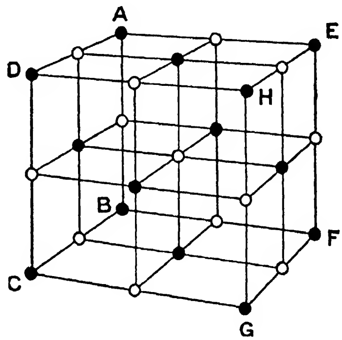


FIG. 403B.

reflections from these two sets of planes will occur at the same setting but the two reflected wave-trains will be out of phase and the resultant intensity will be much reduced. This fluctuation can be observed. With KCl the odd orders practically disappear owing to the matching of the scattering powers of the two kinds of atom and the treatment given above, based on the simple cubic space lattice of Fig. 403A, is justified.

Reference to Fig. 403B will show that the lattice is a repetition of small cubes and each atom lies at the junction of eight cubes. Since also each cube has an atom at each of its eight corners, the whole structure, considered as continued indefinitely in all directions, has one atom, or half a molecule, per cube. Thus if M is the molecular weight of the salt and m is the mass of an atom of hydrogen, namely $1{\cdot}64\times10^{-24}$ gramme, one cube has the mass $\frac{1}{2}Mm$. For NaCl, M=58·5, and each cube has a mass of

$\frac{1}{2}\times 58{\cdot}5\times 1{\cdot}64\times 10^{-24}$ gm. As the density of the crystal is $2{\cdot}17$ gm. cm.$^{-3}$, this mass is also equal to $2{\cdot}17d^3$, whence

$$d=2{\cdot}81\times 10^{-8}\ \text{cm.}$$

Thus for X-rays giving a glancing angle of $11{\cdot}8°$ in the first order, since $\lambda=2d\sin\theta$,

$$\begin{aligned}\lambda&=2\times 2{\cdot}81\times 10^{-8}\times 0{\cdot}204\\&=1{\cdot}15\times 10^{-8}\ \text{cm.}\end{aligned}$$

Reflection by Ruled Grating.—The great penetrating power and the very short wave-length of X-rays make the use of mechanically ruled diffraction gratings impracticable except at nearly grazing incidence. Fortunately then, however, the reflection is very high. Indeed total reflection can be obtained owing to the fact, already mentioned (p. 490), that the refractive indices of many materials are less than unity for X-rays. Moreover, the effective

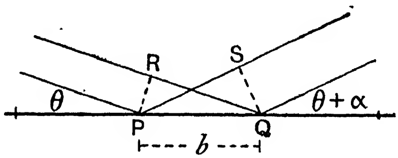


FIG. 404.

spacing between rulings for such extremely oblique incidence is very small and matches the short wave-lengths to be measured.

For incident grazing angle θ (Fig. 404) there will be maximum intensity in the direction $\theta+a$ if the wave-front QS is in advance of PR by a whole number of waves, *i.e.* QR$-$PS$=n\lambda$, where n is integral. Thus

$$\begin{aligned}n\lambda&=b\{\cos\theta-\cos(\theta+a)\}\\&=2b\sin\tfrac{1}{2}(2\theta+a)\sin\tfrac{1}{2}a\end{aligned}$$

and for very small angles,

$$\begin{aligned}n\lambda&=2b(\theta+\tfrac{1}{2}a)\,.\,\tfrac{1}{2}a\\&=b(\theta a+\tfrac{1}{2}a^2),\end{aligned}$$

a quadratic for a with one root given by

$$a+\theta=\sqrt{\left(\theta^2+\frac{2n\lambda}{b}\right)}.$$

The dispersion is

$$\begin{aligned}\frac{da}{d\lambda}&=\tfrac{1}{2}\,.\,\frac{2n}{b}\Big/\sqrt{\left(\theta^2+\frac{2n\lambda}{b}\right)}\\&=\sqrt{\left(\frac{n}{2b\lambda}\right)}\ \text{approximately.}\end{aligned}$$

Unlike the normal optical case, the dispersion depends on the wave-length λ.

The grating space b between successive rulings may be measured with a microscope or found by using light of known wave-length.

This method of measuring X-ray wave-lengths has the great merit over crystal methods of avoiding uncertain assumptions about refraction and about the perfection of the reflecting layers of the crystals used. The values of wave-length found by using gratings tend to be higher (by some 0·1 to 0·3 per cent.) than those found by crystal methods, presumably because the value deduced for d is in fact too small.

X-ray Spectra and Atomic Number.—A comprehensive study of the spectra produced by the reflection of X-rays at the surfaces of a crystal of potassium ferrocyanide was made by Moseley.[1] He used the metal, whose radiations were to be examined, as the target in the X-ray tube, in the manner suggested by Kaye (p. 488). The metals were mounted upon a carrier, which could be moved along so that each metal in turn could be used as target, thus obviating the necessity of exhausting the tube for the employment of each metal. Great care was taken in measuring the glancing angles with accuracy, a spectrometer being modified for the purpose. The reflected rays fell upon a photographic plate calibrated by check observations. Owing to the variable penetrability of the different X-rays used, the aluminium window by which the rays escaped from the tube, when hard rays are examined, was replaced by a window of varnished gold-beater's skin for the examination of the soft rays, and, in addition, the chamber containing the photographic plate was exhausted to reduce absorption by the air.

Characteristic radiations from metals, varying from aluminium to gold, were examined, and the series has since been extended by later workers with results concordant with the earlier results. It was found by Moseley that the spectra from the metals from aluminium to silver consisted of two lines belonging to the series K, the stronger ($K\alpha$) having a greater wave-length than the weaker ($K\beta$), but the lines form a continuous series. On plotting the square root of the frequency of the radiation against the atomic weight of the element, a line, nearly straight, was obtained for each series. But on employing the number of the element in the periodic table, instead of the atomic weight, the lines became almost exactly straight, as shown in Fig. 405. On using the metals zirconium to gold, it was found that the series L radiations each consisted of a series represented by five lines ($L\alpha$, $L\beta$, etc.), but the same linear relation exists. For each series an equation $\nu = a(Z-b)^2$ is found to fit the experimental results, ν being the frequency of the characteristic radiation, and Z a

[1] H. G. J. Moseley, *Phil. Mag.*, **26**, p. 1024, 1913; and **27**, p. 703, 1914.

number which represents the position of the element in the periodic table, and called by Moseley the *atomic number* of the element. There are three exceptions to the position in the periodic table being given by the atomic weight, namely, in the cases of argon, cobalt and tellurium, but the new arrangement places these elements in a position corresponding to their chemical properties.

For the series Kα, $b=1$, and for the series Lα, $b=7{\cdot}4$, which indicates a regular progression in the character of the emission with increase in atomic number. From the value of the constant

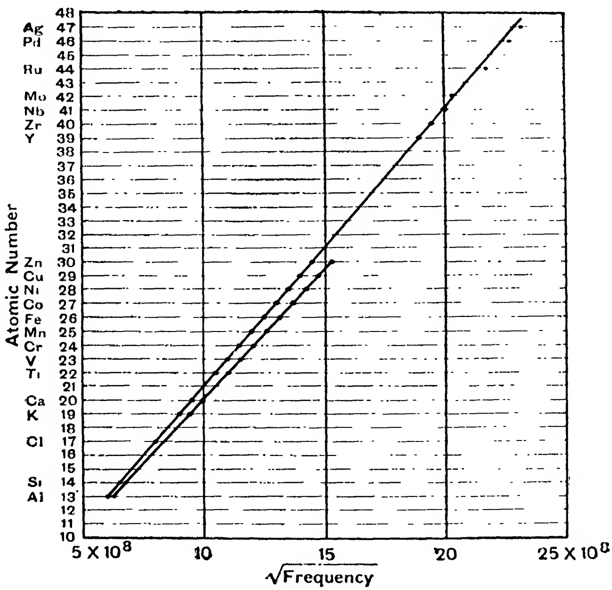


FIG. 405.

a Moseley showed that the results are consistent with the fact, known from other considerations, that Z is the number of positive units of electric charge in the atomic nucleus, the positive unit being the amount which, with one electron, gives a neutral body. Thus the atomic number of hydrogen is 1, helium 2, lithium 3, etc., up to uranium 92.

Canal Rays.—In the cathode dark space, negative ions or corpuscles are driven away from the cathode with a velocity of the order of 10^9 cm. per second (p. 472). If, then, there are positive ions in the cathode dark space we should expect that they would be driven towards the cathode, but since they meet the cathode

itself they would be undetected. The faint glow at the surface of the cathode is due to these positive ions, and, further, these rays carrying a positive charge have been noticed by Perrin (p. 468). If, however, the cathode consist of a thin sheet with perforations in it, the positive ions might then pass through these spaces and give rise to streams behind the cathode. Goldstein [1] observed such rays and called them "Kanalstrahlen," or Canal Rays. They can produce phosphorescence, and are deflected in a magnetic field, but the deflection is much less than in the case of the cathode rays.

Using the method of the combined magnetic and electrostatic fields (p. 471), W. Wien [2] determined the values of $\frac{m}{e}$ and v, and found that $\frac{m}{e}=1\cdot3\times10^{-3}$, and $v=3\cdot6\times10^{7}$ cm. per second. The magnetic deflection is much more difficult to obtain than in the case of the cathode rays, and much stronger fields were used. The velocity is only about $\frac{1}{100}$ of that of the cathode rays, while the value of $\frac{m}{e}$ is of the order of that of the hydrogen atom dealt with in electrolysis. The value of $\frac{m}{e}$ is not so constant as in the case of the cathode rays, but the smallest value found is $1\cdot3\times10^{-3}$.

Positive Ray Analysis.—The streams of positively charged bodies first noticed in the case of the canal rays (above) and afterwards in the α-rays from radioactive substances (p. 516) are usually grouped under the name of *positive rays*, which name indicates the nature of the electric charge carried by them. The positive rays in the discharge tube are of a complex nature. By using large vacuum vessels so that the discharge could be obtained at very high p.d. without injuring the tube, Sir J. J. Thomson [3] found the existence of positive rays whose nature depended upon the gas present in the tube.

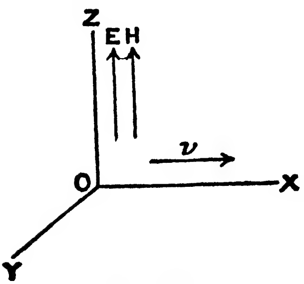


FIG. 406.

The method employed was that of the application of an electric and a magnetic field coincident in position and direction. In Fig. 406 consider the electric and magnetic fields to be both parallel to the axis OZ, and the particle to be moving with

[1] Goldstein, *Berl. Sitz. Ber.*, p. 691. 1886.
[2] W. Wien, *Wied. Ann.*, 65, p. 440. 1898.
[3] J. J. Thomson, *Phil. Mag.*, 21, p. 225. 1911.

velocity v parallel to OX. Then $v=\frac{dx}{dt}$, and the force on the particle due to the magnetic field is Hev (p. 471) and acts in the direction OY. The acceleration of the particle in this direction being $\frac{d^2y}{dt^2}$,

$$m\frac{d^2y}{dt^2}=e\mathrm{H}v$$

$$=e\mathrm{H}\frac{dx}{dt}.$$

Integrating with respect to t, we have,

$$m\frac{dy}{dt}=\int_0^t e\mathrm{H}\frac{dx}{dt}dt=\int_0^x e\mathrm{H}dx$$

But,
$$\frac{dy}{dt}=\frac{dy}{dx}\cdot\frac{dx}{dt}=v\frac{dy}{dx}$$

$$\therefore\ mv\frac{dy}{dx}=\int_0^x e\mathrm{H}dx$$

If y is the displacement after traversing a distance l in the field, so that $x=l$,

$$mvy=\int_0^l\left(\int_0^x e\mathrm{H}dx\right)dx$$

Then integrating by parts,

$$\int_0^l\left(\int_0^x e\mathrm{H}dx\right)dx=\Big[x\Big]_0^l\int_0^l e\mathrm{H}dx-\int_0^l xe\mathrm{H}dx$$

$$=e\int_0^l(l-x)\mathrm{H}dx$$

If now $\mathrm{A}\equiv\int_0^l(l-x)\mathrm{H}dx$, A depends only upon the distribution of the magnetic field and the path of the particles in the field, and is the same for all particles,

$\therefore$
$$mvy=e\mathrm{A}$$

or,
$$y=\frac{e}{m}\cdot\frac{1}{v}\cdot\mathrm{A} \qquad \text{(i)}$$

In order to find the deflection of the particle due to the electric field of strength E, note that the force on the particle is eE.

$$\therefore\ m\frac{d^2z}{dt^2}=e\mathrm{E}$$

Now,
$$\frac{dz}{dt}=\frac{dz}{dx}\cdot\frac{dx}{dt}=v\frac{dz}{dx}$$
$$\frac{d^2z}{dt^2}=v\cdot\frac{d^2z}{dx^2}\cdot\frac{dx}{dt}+\frac{dv}{dt}\cdot\frac{dz}{dx}$$
$$=v^2\frac{d^2z}{dx^2},$$

neglecting the term in $\frac{dv}{dt}$, since the change in velocity is negligible to a first order of accuracy for small displacements,

$$\therefore\ mv^2\frac{d^2z}{dx^2}=e\mathrm{E}$$
$$mv^2\cdot\frac{dz}{dx}=e\int_0^x \mathrm{E}dx$$
$$mv^2z=e\int_0^l\left(\int_0^x \mathrm{E}dx\right)dx$$
$$=e\mathrm{B}$$

where B depends only on the distribution of electric field along the path of the particle, and is the same for all particles.

$$\therefore\ z=\frac{e}{m}\cdot\frac{1}{v^2}\cdot\mathrm{B}\quad\quad\quad \text{(ii)}$$

From equations (i) and (ii)

$$\frac{y}{z}=v\cdot\frac{\mathrm{A}}{\mathrm{B}}\quad\quad\quad \text{(iii)}$$

and,
$$\frac{y^2}{z}=\frac{e}{m}\cdot\frac{\mathrm{A}^2}{\mathrm{B}}\quad\quad\quad \text{(iv)}$$

If a photographic plate be placed to receive the particles, the plate will be affected where the particles strike it. Any particle which is neutral as regards electric charge will not be deviated by the electric and magnetic fields, and will produce a spot upon the plate which marks the original direction of motion of the particles. After exposing and developing the plate it may be placed in the position shown in Fig. 407, so that the axis Oy is parallel to the displacement due to the magnetic field, and Oz parallel to that due to the electric field. A straight line such as OG passing

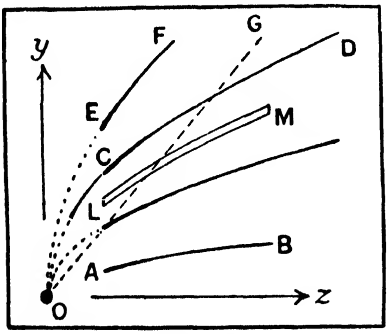


FIG. 407.

through the origin is the locus of the points of striking of all particles having some particular velocity, for $\frac{y}{z}$=constant, and from equation (iii) above this corresponds to a constant value of v, for A and B are the same for all the particles. On the other hand, the points on the plate produced by particles having the same value of $\frac{e}{m}$, will be upon parabolas $\frac{y^2}{z}$=constant, from equation (iv).

The apparatus used by Sir J. J. Thomson is illustrated in Fig. 408. B is a large bulb in which a sufficiently high vacuum to use very high potentials for the production of the positive rays can be maintained. The cathode is an iron tube A having a very fine axial channel, and the electromagnet M supplies the magnetic field. The surface layers of the poles of M are insulated and on being connected to a battery produce the electric field. The fine beam of positive rays subjected to the combined effects of the two fields falls on the photographic plate or other receiving apparatus E.

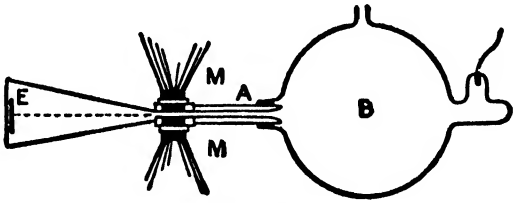


FIG. 408.

Mass Spectra.—Very great increase in the sensitiveness of the method of positive ray analysis, and in the ease of interpreting the results, has been attained by Aston,[1] by means of his mass spectrograph. The positive streams do not pass through the fields simultaneously, but first through the electric field and afterwards through the magnetic field. The former produces a dispersion, which is annulled by the subsequent passage through the magnetic field.

The parallel slits A and B (Fig. 409) reduce the stream of positive rays to a very thin beam which enters the electric field between the plates C and D, slightly inclined to the beam. For simplicity the beam may be considered to undergo a bending θ and a dispersion $\delta\theta$ on arriving at the point E. From equation (ii), p. 500, the displacement of the particle being z, if its length of path in the field is l, then

$$\theta=\frac{z}{l}=\frac{1}{l}\cdot\frac{e}{m}\cdot\frac{1}{v^2}B$$

$$=C\cdot\frac{e}{m}\cdot\frac{1}{v^2}$$

[1] F. W. Aston, *Phil. Mag.*, 38, p. 709. 1919.

where C depends upon the distribution of electric field only, and is the same for all the particles.

Similarly the bending ϕ of the beam in the magnetic field is given by

$$\phi=\frac{1}{l}\cdot\frac{e}{m}\cdot\frac{1}{v}\cdot \mathrm{A}$$

$$=\mathrm{D}\cdot\frac{e}{m}\cdot\frac{1}{v}$$

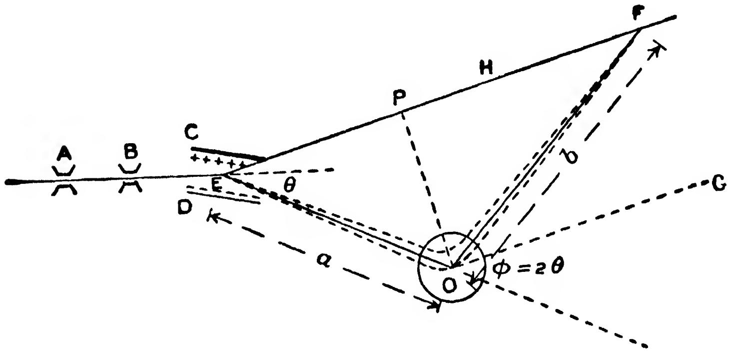


FIG. 409.

where D depends only on the distribution of magnetic field.

Then,
$$\frac{d\theta}{dv}=-2\mathrm{C}\cdot\frac{e}{m}\cdot\frac{1}{v^3}$$

$$=-2\frac{\theta}{v}$$

$$\therefore\ \frac{d\theta}{\theta}+\frac{2dv}{v}=0$$

Similarly,
$$\frac{d\phi}{\phi}+\frac{dv}{v}=0$$

and on combining these two equations we have

$$\frac{d\theta}{\theta}=\frac{2d\phi}{\phi}$$

provided that $\frac{e}{m}$ is constant.

The width of the beam at the point O, whose distance from E is equal to a, is $ad\theta$; and if the beam had travelled a further distance b without meeting the magnetic field the width of the beam would have been $(a+b)d\theta$. But if this divergence is

annulled by the magnetic field, so that the beam comes to a focus at F, where $OF=b$,

$$(a+b)d\theta=b\,.\,d\phi$$

or,
$$\frac{a+b}{b}=\frac{d\phi}{d\theta}=\frac{\phi}{2\theta}$$

and,
$$b(\phi-2\theta)=a\,.\,2\theta$$

On placing the photographic plate in the position HF in the line EPF, the position of the image at F is fixed. EF is parallel to OG, making $\phi=2\theta$.

The image produced in the direction OG would be at infinity, for

$$\frac{b}{a}=\frac{2\theta}{\phi-2\theta}$$

and $b=\infty$ when $\phi=2\theta$.

For $\phi=4\theta$, $b=a$, which would give a focus on HF.

With the small values of ϕ and θ employed, the images for beams of rays having different values of $\frac{e}{m}$ are in focus upon the plate. Also the ratios $\frac{m}{e}$ for the particles are proportional to the distances of the images from some "fiducial spot" which is never far from P.

In practice the deductions are entirely empirical, the positions of the images for substances known to be present serving to fix the values of $\frac{m}{e}$ for the other substances.

In Fig. 410 various mass spectra are shown. The lines for carbon (atomic weight 12), oxygen (16), CO (28), and CO_2 (55) serve to fix the scale for the other lines. The series C (12), CH (13), CH_2 (14), CH_3 (15), and CH_4 (16) are clearly shown in spectrum V. With argon (40) present, the line is clearly marked in VI, and the second and third order lines (A++) and (A+++), that is, the argon atom with two or three electronic positive charges are seen as 20 and 13·3 in V. The atomic weight of neon is 20·20, and with this gas the lines 20 and 22 are obtained (spectrum I), the 20 line being the more intense. This indicates that ordinary neon is a mixture of two constituents of atomic weights, 20 and 22, present in the proportion 9 : 1, which gives the average atomic weight 20·20

Substances such as the two constituents of neon which have identical chemical properties, but different atomic weights, are called *isotopes*. It is now recognised that isotopes have the same atomic number and the same electric charge in the nucleus of the atom (p. 539). The isotopes of any substance, owing to their

identical chemical properties, cannot be separated by any chemical means. Aston has shown that in the case of elements whose atomic weight is a fractional quantity, the ordinary substance is a mixture of two or more isotopes whose atomic weights

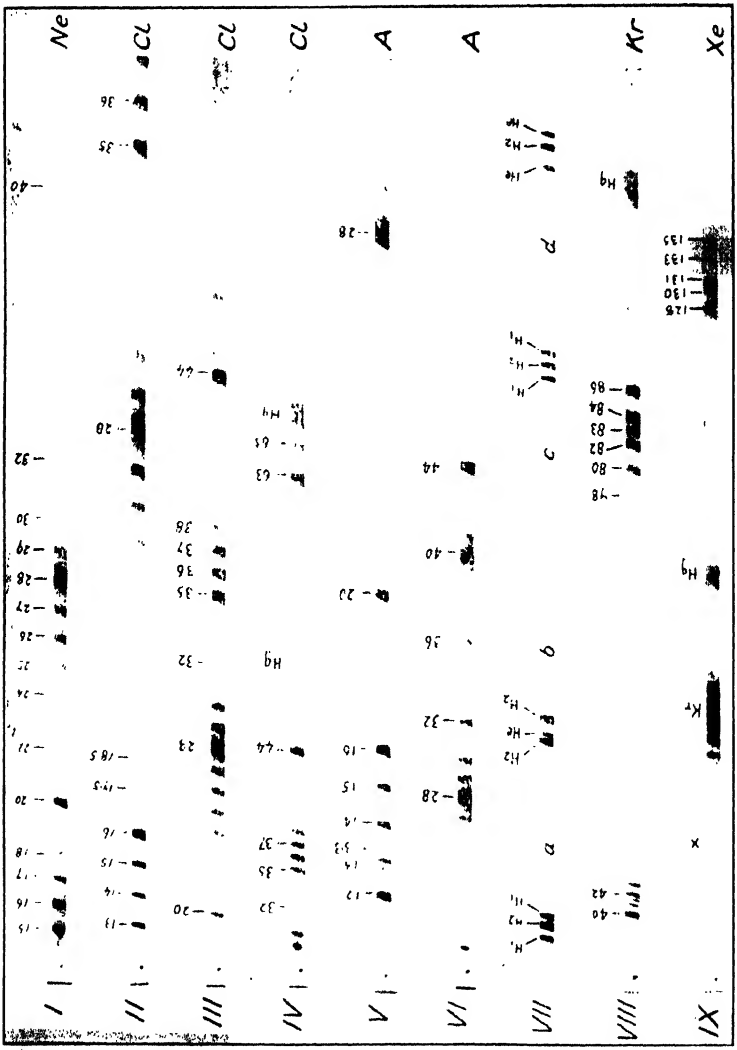


Fig. 410.
(By kind permission of the Editors of the "Philosophical Magazine.")

are whole numbers, that of oxygen being taken as 16. This is very well illustrated in the case of chlorine, whose atomic weight is 35·46. In spectrum III, four lines, 35, 36, 37 and 38, are seen. The lines 36 and 38 correspond to HCl and do not give second order lines, since a molecule does not, as a rule, acquire two

fundamental units of positive charge. The chlorine atoms, however, do acquire two units of charge, as shown by the second-order lines 17·5 and 18·5 in spectrum II. The lines 63 and 65 in spectrum IV are for COCl, and are further evidence of the presence of the two isotopes 35 and 37 of chlorine. Krypton, which has an atomic weight of 83·7, is found to consist of 6 isotopes, 78, 80, 82, 83, 84 and 86 (spectrum VIII). The isotopes of a few of the lighter elements are listed below, the masses being on a scale which assigns the value 16·000 to the commonest oxygen isotope, denoted by the symbol ^{16}O. Z is the atomic number (p. 496). A chart giving in graphical form the well-established isotopes of all the known elements is given at the end of the book (pp. 628–9).

EXAMPLES OF ISOTOPES OCCURRING IN NATURE.

Z=1	Z=2	Z=3
Hydrogen	**Helium**	**Lithium**
$^{1}H=1·0081$ (99·98%)	$^{4}He=4·0039$	$^{7}Li=7·0182$ (92·5%)
$^{2}H=2·0147$ (0·02%)	^{3}He (very rare)	$^{6}Li=6·0169$ (7·5%)

Z=6	Z=8	Z=11
Carbon	**Oxygen**	**Sodium**
$^{12}C=12·0040$ (98·9%)	$^{16}O=16·0000$ (99·76%)	$^{23}Na=22·9961$
$^{13}C=13·0076$ (1·1%)	$^{18}O=18·0037$ (0·20%)	
^{14}C (very rare)	$^{17}O=17·0045$ (0·041%)	

Thus in nature 92·5 per cent. of all lithium consists of atoms with mass close to 7, the remaining atoms having the mass 6. All the atomic masses on this scale are very nearly integral. About 30 per cent. of the elements up to bismuth (Z=83) have only one known stable isotope, but many elements have several isotopes. Tin has no less than 10 (masses 112, 114, 115, 116, 117, 118, 119, 120, 122 and 124).

Deuterium, the Isotope of Hydrogen.—After allowance has been made for the presence in water of oxygen atoms of masses 17 and 18, the ratio of the mass of hydrogen to that of oxygen in a sample of water betrays the presence of a hydrogen isotope of double the normal mass. The presence of this isotope in hydrogen gas should give rise to faint lines accompanying the ordinary lines of the spectrum but displaced to the red (long-wave) side. These were sought and detected by Urey, Brickwedde and Murphy.[1] It was then found by Washburn and Urey[2] that water in old electrolytic cells contained a high percentage of heavy water. This arises because of the lower velocity of migration (p. 177) of

[1] H. C. Urey, F. G. Brickwedde and G. M. Murphy, *Phys. Rev.*, **39**, p. 164.
[2] E. W. Washburn and H. C. Urey, *Proc. U.S. Nat. Acad. Sci.*, July, 1932.

the more massive ions containing the "heavy" isotope, to which the name *deuterium*, symbol D, has been given. Thus pure heavy water is D_2O. Production is chiefly by electrolysis of sodium hydroxide with nickel electrodes.

Heavy water is not radically different in physical properties from ordinary water, but has a higher freezing point (3·8° C.), boiling point (101·42° C.), viscosity and density (1·104). It has a maximum density at 11·6° C.

The ordinary charged hydrogen ion of the mass spectrograph is a hydrogen nucleus or *proton*; the corresponding particle of double mass is called a *deuteron*.

Ionisation by Collision.—Many facts point to the conclusion that the ions which take part in the passage of currents through gases are to a large extent produced by the impact of ions already present, with the neutral atoms of the gas. An examination of the curve in Fig. 390B shows that when a certain electrical intensity of the field is reached, a large and rapid increase in the current takes place, and it is reasonable to suppose that this happens when the velocity of the ions due to the electric field is sufficient for them to ionise the neutral molecules of the gas on impact. The conditions for this to take place are complicated; at high pressures the collisions are so frequent that the ion will not have acquired a sufficient velocity before impact to enable it to produce ionisation, but on the other hand, if the potential gradient be very great, a much shorter path is required for this critical velocity to be produced. This is in accordance with the fact that at high pressures a much greater potential difference is required to produce a spark than at low pressures, the length of spark gap remaining the same.

If l be the mean length of path of the ion between collisions,

$$e\mathrm{E}l=\tfrac{1}{2}m\,.\,v^2,$$

since the work done on the ion by the electrical intensity E is equal to the kinetic energy acquired by the ion. We therefore see that the negative ion, having a much smaller mass than the positive ion, will acquire the ionising velocity in a much shorter path than the heavy positive ion, and therefore at the beginning of the discharge the negative ions will be the more important in producing ionisation. But the phenomenon is complicated by the fact that the positive ion may not require the same velocity to produce ionisation as the negative ion, and, further, the collisions do not take place under the same conditions.

Process of Electric Discharge.—The process of the electric discharge may now be accounted for on the theory of ionisation by collision. When there is an electric field in the gas between two conductors, the current will be infinitesimal (although never

actually zero), unless ions are produced by some external cause, such as Röntgen rays or ultra-violet light. When, however, the electric intensity reaches a certain value, any ions in the gas will acquire a velocity sufficient to produce ionisation by collision. A few ions are always present, for no gas is a perfect insulator. C. T. R. Wilson[1] found that at the atmospheric pressure, the rate of leakage of charge from a body in an enclosed space is $10^{-8}v$ electrostatic units per second, where v is the volume in cubic centimetres of the enclosure. When the electric intensity reaches such a value that ionisation by collision begins, the number of ions present will rapidly increase, and it is found that a very much smaller electric intensity is required to maintain the current in the gas than to start it.

Prof. Townsend[2] found in one case, with air at 4·31 mm. pressure, between parallel electrodes 8 mm. apart, that the gas acted as an insulator when the difference of potential between the plates was 601 volts, but on increasing this to 603 volts a current of 0·0052 ampere passed between the electrodes, the difference of potential between which dropped to 350 volts.

The lag in the establishment of the spark that has been noticed by many observers is also accounted for, the gas acting as insulator, for a very short time, to an electromotive force which would produce the discharge if continuously applied. The setting up of the steady condition of ionisation requires time, since the initial number of ions present in the gas is exceedingly small.

The high-voltage spark and the lightning discharge have been extensively studied in recent years. The actual form of the discharge path may be photographed by a device due to C. V. Boys[3] in which rapidly moving lenses and a stationary film are used, or stationary lenses and a film secured to a drum rotating at high speed. By this and other means it has been shown that the main discharge is preceded in general by a "leader stroke" which makes progress in a series of steps, heavy ionisation of the gas being built up until a conducting path is formed, whereupon the main discharge passes. The phenomena are clearly very complicated. The interested reader may consult an article on the lightning discharge by J. M. Meek and F. R. Perry.[4]

Cosmic Rays.—The conductivity of the atmosphere, which was at first attributed to radioactive products diffused in the air, took on a different aspect when it was discovered by Gockel[5] that at a height of 4500 metres the conductivity is greater than at the earth's surface. Subsequent investigation established that the

[1] C. T. R. Wilson, *Proc. Roy. Soc.*, **68**, p. 151. 1901
[2] J. S. Townsend, *Phil. Mag.*, **8**, p. 738. 1904.
[3] C. V. Boys, *Nature*, **118**, p. 749 (1926) and **124**, p. 54 (1929).
[4] *Reports on Progress in Physics*, Vol. X. The Physical Society, 1946.
[5] A. Gockel, *Phys. Zeit.*, XI, p. 28. 1910.

conductivity increases with altitude and that there is little difference in the values between day and night. There is now no doubt that this conductivity is mainly due to exceedingly penetrating radiation coming from outer space and accordingly called *cosmic rays*.

Using ionisation chambers to measure the intensity of the cosmic radiation, Otis, Cameron and Millikan determined the absorption coefficient for cosmic rays in the water of high-altitude lakes, finding the rays to be 18 times as penetrating as the most penetrating terrestrial source available to them, the γ-rays emitted by radioactive substances, described in the next chapter. They immersed electroscopes to various depths in the lakes and they also made measurements at high altitudes by means of balloon-borne recording electrometers. They concluded that the whole atmosphere, regarded as an absorber, is equivalent to about 10 metres of water, which is the height of the water barometer, and also that the radiation contains constituents of different degrees of penetrability.

The incidence of the rays is not quite uniform all over the earth, and the variations are connected with the terrestrial magnetic field. Thus there is a reduction in intensity (of the order 10 per cent.) on passing from high magnetic latitudes to the vicinity of the magnetic equator, a phenomenon known as the *latitude effect*. This is to be expected if the radiation contains electrically charged particles, for a charged particle approaching the earth experiences a transverse force which depends on the component of the magnetic field at right angles to its path (cf. pp. 240 and 471). This component is, for rays directed more or less vertically down, greatest near the magnetic equator. Unless the incident energy of such particles is very high, they may be deflected away from the earth altogether.

The same considerations show that there should be an *East–West effect*, a preponderance of particles arriving from one side or the other, depending on their sign. To investigate this, use has been made of the Geiger tube discussed more fully in the next chapter. This is in essence an ionisation chamber which is used very near the limit of stability so that a very little extra ionisation will " trigger off " a discharge (cf. p. 507). The surge of charge so produced gives an audible signal (by amplification and use of a loud-speaker) or it can be made to operate a counter. If two such tubes are so connected together that discharges and therefore signals are only given when simultaneous ionisations occur, the device is virtually responsive only to radiation coming along, or very close to, the direction defined by the line passing through the centres of the two tubes. This direction may be called the " axis " of the " cosmic-ray telescope." By using pairs of tubes in this

way, and comparing the rates of receipt of signals from units with axes pointing in different directions, the intensities associated with different directions of arrival may be deduced. There is found to be a preponderance of arrivals from the westerly side of the local magnetic meridian, which implies, as the reader may verify, that the incoming radiation contains predominantly positively charged particles.

The ionising radiation at ground level contains many fast particles and these may give tracks in a Wilson Cloud Chamber (p. 482) if an expansion is made very soon after such a particle has passed through the chamber. Condensation of minute water droplets gives under low magnification the appearance of a continuous trace. Blackett and Occhialini designed a cloud chamber arrangement in which an expansion automatically occurred whenever a signal was received by a pair of Geiger tubes connected so that the axis of the pair passed through the chamber. In this device, a photographic exposure is also automatically made and in this way great numbers of tracks of particles have been recorded and examined.

The tracks in some of the photographs are forked, indicating that a violent collision has occurred with a gas molecule in the chamber, both the original particle and the struck molecule giving tracks. A collision of this kind is clearly quite different from the less violent impacts which the particle must be supposed to be making in great number along its track, to account for the trail of ionised molecules on which condensation occurs. As will appear later, an atom contains an extremely small charged nucleus, accounting for nearly all its mass, and the forked tracks denote a more or less direct hit on a nucleus.

The Positron.—In order to study the energy of the electrons liberated by cosmic rays, Millikan and Anderson in 1929 constructed a vertical closed chamber of the Wilson type (p. 482) of very large dimensions. By means of an electromagnet a strong magnetic field could be produced. In this way it was possible to produce a nearly uniform field over an area of 17 cm. $\times$17 cm. Since the photographs of the cloud tracks are taken in a plane at right angles to the magnetic field, the radius of curvature, r, of the track will be given by equation $\mathrm{H}r=\frac{mv}{e}$ (p. 471) where m, v and e refer to the particle causing the track. The kinetic energy $\frac{1}{2}mv^2$ of the particle may be expressed in electron-volts, V, where $\frac{1}{2}mv^2=e_1\mathrm{V}$, e_1 being the electronic charge. Then, $\mathrm{H}r=\frac{m}{e}\sqrt{\frac{2e_1\mathrm{V}}{m}}=\sqrt{\frac{2me_1\mathrm{V}}{e^2}}$. Thus if the mass and charge of a particle are known, its energy in electron-volts may be found

from the strength of magnetic field and the radius of curvature of the cloud track. It was found that when the γ-rays from Th C″ of maximum energy $2{\cdot}6 \times 10^6$ e.v. are passed through the expansion chamber, the electrons liberated were caused to move in spirals whose radius indicated an energy of 2×10^6 e.v. On the other hand, cosmic rays liberated, in one case, an electron of energy 8×10^6 e.v. At this time many instances occurred of tracks appearing in pairs, one having its curvature in such a direction that the electric charge of the particle must be positive. At first these positive particles were considered to be protons. This, however, was definitely ruled out by a photograph obtained by Anderson[1] in which one of these bodies traversed a plate of lead 0·6 cm. thick. The curvature must be greater where the velocity is less, which, as the curvature was greater on one side of the plate than on the other, showed the direction of travel. Since the charge was positive and the ionisation of the type produced by electrons, it was concluded that the body was a positive charge of mass equal to that of the electron. It has been named the *positron*. Amongst the high energy rays it is now known that positrons are as common as electrons.

Amongst others, Chadwick, Blackett and Occhialini[2] showed that when the rays from a mixture of polonium and beryllium fall on a lead plate in an expansion chamber positrons are produced whose direction is fixed by the loss in energy in traversing the lead plate.

Positrons are also emitted by substances which exhibit induced radioactivity.

Cosmic Ray Showers.—By employing the large cloud chamber (p. 509) it has been found that two or more tracks starting from the same point are of frequent occurrence. These associated tracks constitute a *shower*, and consist of both electrons and positrons. The chance of obtaining a shower in the chamber is increased by placing a lead sheet in the chamber. The device of the pair of Geiger tubes, already described on p. 509, is used to select many-track showers. There may be many hundreds of particles in a shower and the total energy of the whole may exceed 2×10^9 e.v.

In many of the photographs small circular tracks appear. The strong curvature in the magnetic field shows that these are due to charged particles of comparatively low energy. They are presumed to be liberated by radiation emanating from atoms of the gas on being struck by other particles in much the same way as X-radiation is emitted when the atoms of the target of an X-ray tube are struck by electrons.

[1] C. D. Anderson, *Science*, LXXVI, p. 238. 1932.
[2] J. Chadwick, P. M. S. Blackett and G. Occhialini, *Nature*, **131**, p. 473. 1933.

Origin of the Cosmic Rays.—There has been much speculation about the origin of the cosmic rays. It was held for a long time that the primary radiation must be photons, *i.e.* radiation, because the most likely alternative appeared to be electrons or protons and it seemed difficult to explain how such particles could carry the immense concentration of energy displayed in a shower.

R. A. Millikan suggested that the cosmic rays were radiation created by the annihilation of matter. According to Einstein, mass is a form of energy, and the energy-equivalent of a mass m gm. is mc^2 ergs, where $c=3\times10^{10}$ cm. per sec., the velocity of light in a vacuum. An electron of charge e electromagnetic units, accelerated by a potential difference of V volts, acquires energy $e\mathrm{V}\times10^8$ ergs, and equating this to mc^2 to find the energy-equivalent of the electronic mass in electron-volts,

$$\mathrm{V}=\frac{c^2}{\dfrac{e}{m}\times10^8}=\frac{9\times10^{20}}{1{\cdot}757\times10^7\times10^8}$$

$$=512{,}000 \text{ volts.}$$

Annihilation of an atom of hydrogen would yield 1835 times as much, or $9{\cdot}397\times10^8$ e.v. Dividing by 1·0078, the atomic weight of hydrogen, a fall of 1 in an atomic weight would liberate $9{\cdot}32\times10^8$ e.v. per atom affected. Union of four hydrogen nuclei to form a helium nucleus would cause a loss of atomic weight, known as the *packing effect*, of some $(1{\cdot}008\times4-4{\cdot}004)=0{\cdot}028$, which is equivalent to some $3{\cdot}2\times10^6$ e.v. Many other "nuclear reactions" can be conceived which involve higher energies.

Nuclear reactions are believed to occur in the interiors of the sun and other stars and to account for the enormous output of radiated energy from these bodies, but it is not now thought that photons liberated in such transmutations constitute the primary cosmic radiation. Experimental study of the phenomena at high altitudes, discussed at the end of Chapter XVI, shows that the primary radiation is almost certainly corpuscular, consisting of relatively massive atoms or ions moving at very high speed. Their origin is still somewhat obscure.

CHAPTER XV

RADIOACTIVITY

Becquerel Rays.—While investigating the relation between phosphorescence and the production of rays able to produce a photographic effect after transmission through opaque material, Becquerel[1] found, that in the case of the double sulphate of uranium and potassium, a photographic effect was produced even when the salt had not been exposed to sunlight. In fact, it was subsequently found that the effect is the same after keeping the salt in a light-tight lead box for several years, or on dissolving it in water in the dark and recrystallising it, still in the dark ; and further, that the photographic effect is produced by the uranium, whatever the nature of the chemical combination in which it exists.

The photographic effect is of a similar nature to that produced by the Röntgen rays, but is very much feebler than the effect produced by an ordinary X-ray tube. Whereas an exposure of a few seconds to Röntgen rays will produce a considerable photographic effect on an ordinary sensitive plate, several days' exposure is necessary in the case of uranium.

Other substances have been found to emit rays similar to those emitted by uranium, and the name of Becquerel rays has been given to them ; but, owing to the complexity of these rays, other names for the several constituents have replaced the original name for general use.

Ionisation.—The Becquerel rays possess the power of rendering the gas through which they pass conducting. Thus if the uranium salt be spread upon the plate A (Fig. 411) parallel to the plate B, which latter is in connection with the gold leaf of an electroscope, the charge given to the electroscope will leak away, owing to the gas between A and B being a conductor, and the rate of leak is a measure of the ionising power of the substance spread upon A. This form of the electroscope was used by M. and Mme. Curie in many of their investigations.

Another extremely useful form of electroscope for the measurement of the ionisation produced by radioactive substances is due to C. T. R. Wilson. The rigid support B has a thin aluminium

[1] H. Becquerel, *Comptes Rendus*, **122**, p. 501. 1896.

leaf A attached to it. When charged, the leaf stands as shown in Fig. 412, and its motion, as observed by a telescope with a transparent scale in the eyepiece, or by comparison with the reflected image of a linear scale, is a measure of the conductivity of the gas within the cubical brass vessel. In order to obtain good insulation, B ends in a metal block, C, carried by a tube of fused quartz, D, fixed by shellac to the brass tube E. This is supported by an ebonite bush, F, in the upper and lower faces of which annular slots are turned and filled with sulphur to prevent

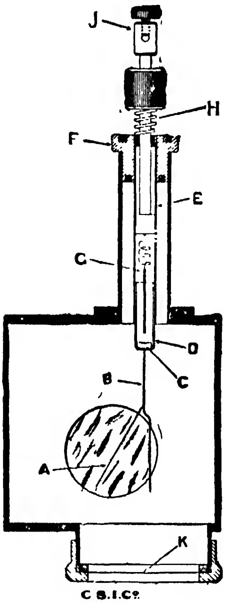


Fig. 412.

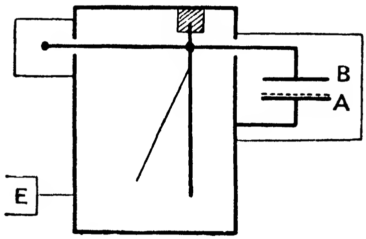


Fig. 411.

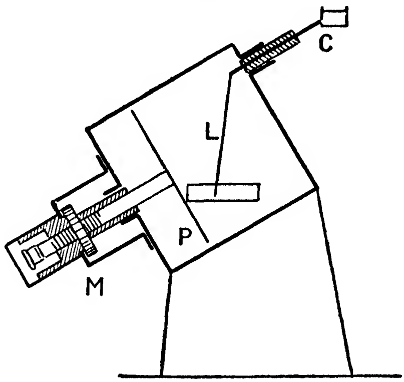


Fig. 413.

leakage over the faces. In order to charge the leaf, the brass rod which carries the terminal J at its upper end and the light rod G at its lower end, is depressed until G touches C, and the charge is then given to the leaf. On releasing the rod, the spring H raises it, and B and A are again insulated. The lower window K is covered with a layer of thin tissue paper to exclude draughts, and the radioactive material is placed below it, the rays which produce ionisation thus entering the chamber of the electroscope.

A more sensitive arrangement has also been devised by C. T. R. Wilson,[1] in which the gold leaf L (Fig. 413) is attracted by the plate P, which is charged to a constant potential of about 200

[1] C. T. R. Wilson, *Camb. Phil. Soc. Proc.*, **12**, p. 135. 1903

volts. The best form of the instrument, and the conditions for satisfactory working, have been found by G. W. C. Kaye.[1] The gold leaf is first connected to the brass case, and the instrument tilted until the leaf is in the field of the observing microscope. The sensitiveness of the instrument can be altered by varying its tilt, and also the distance of the earthed plate P from the gold leaf, by means of the micrometer screw M, the maximum sensitiveness occurring when the leaf approaches instability owing to its proximity to the plate. The leaf is then connected by means of the conductor C to the body whose rate of change of potential it is required to know. With the plate at potential 207 volts, and the leaf inclined at 30°, a travel of about $5\frac{1}{2}$ mm. was found for a variation in potential of the leaf of 1 volt.

The quadrant electrometer also is extensively used for the measurement of the current in ionised gases. The radioactive material is spread upon the plate A (Fig. 414), the parallel plate

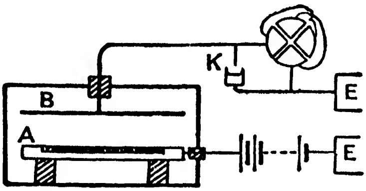


FIG. 414.

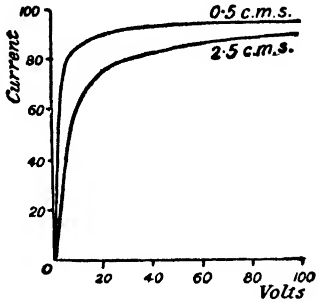


FIG. 415.

B being connected to one pair of quadrants of the electrometer. One end of a battery is connected to A, the other end being earthed. To begin with, the plate B is also connected to earth by means of the key K. This is opened at a known instant, and the electrometer deflections observed after equal intervals of time. Then, as on p. 480, if c is the capacity of the electrometer and the conductors connected with it, and θ the deflection at any time t—

$$i = ck \cdot \frac{d\theta}{dt}$$

where k is the difference of potential between the quadrants for unit deflection.

By varying the electromotive force of the battery used to produce the current, the relation between potential difference and current can be obtained. This relation is similar to that obtained in the case of the conductivity produced by X-rays; that

[1] G. W. C. Kaye, *Proc. Phys. Soc.*, **28**, p. 209. 1911.

is, the current increases rapidly with the difference of potential for small values but soon ceases to increase, the greatest value being the saturation current (Fig. 390B). This depends upon the pressure, potential gradient and the distance apart of the plates and the amount of radioactive material present. The curves, Fig. 415, given by Lord Rutherford show the relative increase in the current with potential gradient for a thin layer of uranium oxide upon one of a pair of parallel plates, when the distance between the plates is 0·5 cm. and 2·5 cm. respectively.

Thorium.—On searching amongst the other elements for the emission of Becquerel rays, it was found by Schmidt [1] that in the case of thorium the emission was about as strong as in that of uranium. Thorium is largely used in the manufacture of the Welsbach incandescent gas mantles, and on laying one of these mantles flat on a photographic plate for about a week, it is found when the plate is developed that the woven pattern of the mantle is seen upon it.

Radium.—On examining a number of minerals for the emission of Becquerel rays, or for radioactivity, as it is now called, Mme. Curie found, using the leakage method, certain specimens of pitchblende to be more radioactive than uranium. The mineral pitchblende contains barium, and on separating out this substance by precipitation as the carbonate, it is found that the radioactivity of the precipitate is very great. On converting into the chloride and employing the method of fractional crystallisation, the parts that separated out first were found to be, as the process was repeated, more and more radioactive. M. and Mme. Curie in this way separated from the barium another substance of enormous radioactivity which they called *radium*. The process of separation of the radium chloride from the ore is exceedingly tedious, a ton of ore yielding only a few decigrammes of radium.

Polonium.—One of the processes in the separation of the metals in the pitchblende consists in the precipitation of the lead, antimony, bismuth group by means of sulphuretted hydrogen. The deposit produced is found to be radioactive, and a further separation showed that the radioactivity is associated with the bismuth. By fractional precipitation by diluting a solution of the nitrate, a new radioactive element which Mme. Curie named *polonium* was obtained. The preparation of a specimen of reasonable purity is very difficult. It has been found[2] to have a melting point of 246° C. and a density between 9 and 10 gm. per c.c. Polonium is Ra F, one of the chain of disintegration products of radium (p. 538).

Actinium.—Another radioactive material has been obtained

[1] G. Schmidt, *Wied. Ann.*, **65**, p. 141. 1898.
[2] C. R. Maxwell, *J. Chem. Phys.*, **17**, p. 1288. 1949.

by Debierne [1] from pitchblende in association with the iron group, and with difficulty separated out. It has been named *actinium*, and has an activity comparable with that of radium.

Absorption.—*α, β, γ Rays.* If a layer of radium bromide be placed in the tray (Fig. 411), and the rate of collapse of the leaves observed, it will be found on covering the radium with a sheet of tinfoil, that the rate of collapse of the leaves is very much less than without the tinfoil. The ionisation may be reduced to one-tenth by a sheet of ordinary foil. If the rays emitted by the radium are all of one kind, a second layer of tinfoil would produce a further proportionate reduction, and the radiation transmitted would be one-hundredth of the original amount. This, however, is not found to be the case; the reduction produced by the second layer of foil is very small. Hence, there are at least two constituents in the original rays, one readily absorbable, and the other much less absorbable. Rutherford named the more absorbable rays the *α* rays and the more penetrating the *β* rays.

On continuing the above experiment with more layers of tinfoil, it will be found that after a time the additional layers again produce less effect; or if sheets of lead be used, it is found that a sheet 2 mm. thick produces a large reduction in the ionisation, but a second sheet of the same thickness does not produce nearly so great a reduction as the first. This is due to the fact that in addition to the *α* and *β* rays, others of very much greater penetrating power are present, which Rutherford called the *γ* rays. The following table is given by him :—

Rays.	Thickness of aluminium which reduces ionisation to one-half.	Relative penetrating power.
α	0·0005 cm.	1
β	0·05 cm.	100
γ	8 cm.	10,000

***α* Rays.**—The ionisation produced by the rays emitted by radium is chiefly due to the *α* rays, owing to the great quantity emitted; the ionisation due to the *β* and *γ* rays has been noted above.

Another property of the *α* rays is their power of producing fluorescence; a diamond exhibits a blue fluorescence when brought near a small quantity of radium bromide. Other substances, such as zinc sulphide, also are caused to fluoresce by the *α* rays, and hence the spinthariscope of Sir Wm. Crookes, in which a speck of radium bromide is placed behind a screen on

[1] A. Debierne, *Comptes Rendus*, 130, p. 906. 1900.

which is spread a thin layer of zinc sulphide, the whole being mounted in a brass tube, at the other end of which is a lens, placed so that an enlarged image of the screen can be seen. It is then observed that the fluorescence is not a uniform glow, but has the appearance of a shower of sparks, no two following each other in the same place. That the fluorescence is due to the α and not the β or γ rays may be proved by interposing a thin sheet of mica between the radium and the screen, the fluorescence then ceasing. The cause of the luminosity is probably the rupturing of the crystals of zinc sulphide when struck by an α ray particle, as a similar luminosity may be produced by fracturing the crystals by mechanical means.

Deflection of α Rays by Magnetic Field.—The α rays require a strong magnetic field for their deflection, but the fact that they can be deflected shows that they consist of moving charged particles, and further, the direction of the deflection proves them to be positively charged. Owing to the small amount of deflection, the method of p. 471 is not applicable. Lord Rutherford[1] measured the deviation in a magnetic field by the method illustrated in Fig. 416. The radium is spread in a thin layer underneath a system of parallel plates, P, placed vertically and at known distances apart. With no magnetic field, the α rays pass vertically upwards between the plates, and passing through the extremely thin aluminium window A, enter the electroscope chamber C and cause a collapse of the leaves at a rate which can be measured. On applying a magnetic field which is horizontal and parallel to the plane of the plate P, the α rays are deviated in such a way that they are driven against the plates, and will not then reach the chamber C. In the left-hand part of Fig. 416 the α rays are shown passing upwards as they do without the magnetic field being present, and those shown on the right hand are being deviated by the field. The field which just cuts off the rays from the chamber C is found, and then the dimensions of the spaces being known, Hr (see p. 471) is known. During the experiment a stream of gas passes downwards through the apparatus to carry away the emanation as it is formed (p. 531).

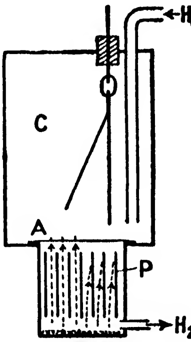


FIG. 416.

The electrostatic deviation of the α rays was found by means of an experiment similar to the above, but with alternate plates connected together, the two sets being maintained at different potentials. The electrical field between the plates caused the

[1] E. Rutherford, *Phil. Mag.*, 5, p. 177. 1903.

α rays to be driven against one set of plates as before, with consequent reduction in the rate of ionisation in C.

By reasoning similar to that on p. 471 it was found that for the α rays from radium—

$$v = 2{\cdot}5 \times 10^9 \text{ cm. per sec.}$$

and,
$$\frac{e}{m} = 6 \times 10^3,$$

which is a quantity of the order of half that for the hydrogen ion in electrolysis. By arranging that the plates P have a projecting ridge on one side at the upper edges in some of the experiments, so that the rays when deflected to this side were not allowed to pass, it was shown that the charges of the α particles are of positive sign. The value of e/m now accepted is $4{\cdot}78 \times 10^3$.

Absorption of α Rays.—Measurements of the absorption of α rays have brought to light the interesting fact that the power of producing ionisation possessed by them does not diminish gradually as their path in the absorbing medium increases; it does not diminish at all up to a certain range, and then ceases abruptly. This discovery was made by Bragg and Kleeman,[1] who used a small quantity of radioactive substance at A (Fig. 417), and limited the α rays to a narrow beam falling upon the

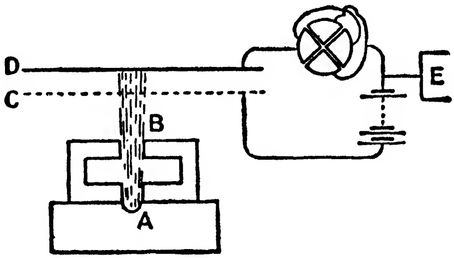


FIG. 417.

air situated between the gauze C and the metallic plate D. These are kept at constant distance apart, their distance from A being variable. The ionisation at any given distance is then measured by the rate of leak of charge between C and D when a constant difference of potential is maintained between them. The "range" of the α rays in air, that is, the distance travelled before their ionising power ceases, is then found by varying the distance of CD from A, until the rate of leak of charge is independent of the presence of the radioactive material.

The layer of radioactive material at A must be very thin, or some of the α rays will on emergence have already passed through

[1] W. H. Bragg and R. Kleeman, *Phil. Mag.*, **10**, p. 318. 1905.

a layer of the material, and their "range" have been reduced, so that the beam is no longer homogeneous and the ceasing of the ionisation will not take place abruptly. The "range" for the α rays produced by radium is in air about $3\frac{1}{2}$ cm., and it is important to notice that the phosphorescent and photographic effect of the rays ceases at the same distance as their power of producing ionisation. Rutherford, by the method of magnetic deflection, found that the ionisation ceases when the velocity of the α particles falls below $1{\cdot}12 \times 10^9$ cm. per second. Hence α rays with velocity less than this would be undetectable except for their charge, and thus many substances may be emitting α rays whose radioactivity is at present unsuspected.

By interposing layers of different materials in the path of the α rays, their effect upon the "range" was found, and the results showed that the stopping power of any substance is proportional to the square root of its molecular weight.

Bragg and Kleeman also found that the α particle spends its energy in producing ionisation at a rate which is proportional to $1/\sqrt{V}$, where V is its velocity, until near the end of its path, when the drop in ionisation is rapid. The curve connecting ionisation and range is of the type shown in Fig. 418, in which the saturation current due to the ionisation produced by the α rays at different distances from the radiating material is shown. They also found that the curve representing the end of the range is identical for the α rays from all substances. If the curve ABCD is that for an α particle of range OD, then the curve BCD is that for an α particle of range ED.

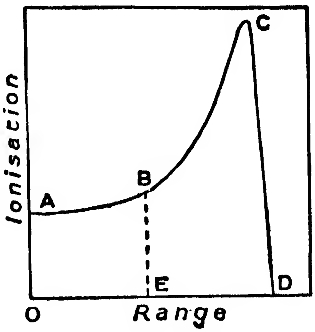


FIG. 418.

Range and Velocity of α Rays.—Geiger[1] found that a very important relation exists between the range of an α particle and its velocity of emission—that the range is proportional to the cube of the velocity, or,

$$R = aV^3.$$

Taking the range in air at 15° C. and 76 cm., pressure to be 6·9 in the case of the α particles from Ra C′ (p. 538), and the velocity to be $1{\cdot}92 \times 10^9$ cm. per sec., it is then possible to calculate the velocity of emission for the α particles emitted by any of the other substances, using the value for the range given in the table on p. 538.

[1] H. Geiger, *Roy. Soc. Proc.*, A, lxxxii, 5. 1910.

β **Rays.**—The fluorescence produced by the β and γ rays is brilliant in the case of barium platinocyanide, which is therefore a convenient substance for studying these rays. Many other substances exhibit fluorescence, the colour varying with the substance.

The absorption of the β rays by ordinary matter has been described on p. 516. They are very much more deviated in a magnetic field than the α rays, and in a direction which indicates that they are negatively charged particles. The beam of β rays from a specimen of radium bromide is not deviated uniformly in a magnetic field, but is spread out, indicating that the beam itself consists of particles in different conditions. Becquerel made measurements upon the magnetic deviations in a number of cases, but the greatest interest attaches to some measurements of Kaufmann,[1] in which the velocity of the rays and the ratio $\frac{e}{m}$ are obtained by causing the displacement produced by a magnetic field, and one by an electrostatic field, to take place simultaneously, but in directions at right angles to each other. A thin beam of β rays falls normally upon the photographic plate, giving rise to a small patch when there is no magnetic or electrostatic field. The magnetic field alone, being at right angles to the rays, would spread them out into a " spectrum " in a line at right angles to the direction of the field. The electrostatic field is in the same direction as the magnetic field, but since it produces a deflection in its own direction, this is perpendicular to that produced by the magnetic field. The method is similar to that of crossed spectra used in optics, and has already been described in connection with positive ray analysis (p. 498). In this way it was found that the ions have velocities much greater than those in the cathode rays, but that the mass varies with the velocity, increasing as the velocity approaches the velocity of light, as the following table shows :—

Velocity.	$\frac{e}{m}$.
$2{\cdot}36\times10^{10}$ cm. per sec.	$1{\cdot}31\times10^{7}$
2·48 „ „ „	1·17 „
2·59 „ „ „	0·97 „
2·72 „ „ „	0·77 „
2·85 „ „ „	0·63 „

The diminution of $\frac{e}{m}$, due to the increase in m in the ratio

[1] Kaufmann, *Phys. Zeitschr.*, **4**, No. 1*b* 1902.

$(1-v^2/c^2)^{-\frac{1}{2}}$ when the velocity increases, is a consequence of the electromagnetic theory and of the theory of relativity, as we shall see on p. 546, where v is the velocity of the particle and c that of light.

The distribution of velocity among the β particles emitted by a radioactive substance has also been investigated in an interesting manner by Rutherford and Robinson.[1] The source of β rays is placed at A (Fig. 419) under a slit S, and the rays are bent into circles ABC, ADC, by means of a magnetic field at right angles to the plane of the diagram, and fall upon a photographic plate, the vessel being exhausted. For β ray particles of a particular velocity, the circles have radii given by $Her=mv$ (p. 471). Thus β particles having the same velocity will move in circles which intersect at C. For a wide slit, a given velocity of ray corresponds to a line formed at C. For the mean circle ASGC, ASC is a right angle, and $AC=2r$, the diameter of the circle. If $AS=2y$ and $SC=2x$, $r^2=x^2+y^2$. The β rays from the

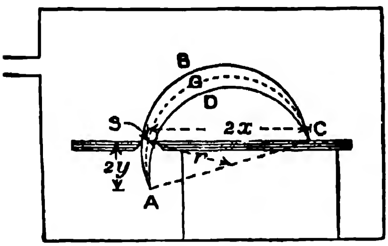


FIG. 419.

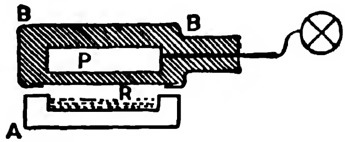


FIG. 420.

material at A are therefore grouped in velocities on the plate, which velocities can be found. Rutherford and Robinson found the β ray spectra so obtained to be complicated. They recognised 16 groups for the β rays from Ra B, and 48 in the case of Ra C. The velocities of the former vary between 0·365 and 0·823 of the velocity of light, and in the case of the latter between 0·632 and 0·986. It was also observed that the energy of each line is nearly an integral multiple of the common difference in the energies of consecutive lines.

Charge carried by β Rays.—That the β rays carry a negative charge has been shown by many experimenters; notably by M. and Mme. Curie,[2] who allowed the rays to fall on a plate connected to an electrometer, and observed the changing deflection. The chief difficulty arises from the fact that the rays render the air surrounding the body conducting, and they got over this difficulty by embedding the conductor P (Fig. 420), which absorbs

[1] E. Rutherford and H. Robinson, *Phil. Mag.*, 26, p. 717. 1913.
[2] M and Mme. Curie, *Comptes Rendus*, 130, 647. 1900.

the rays, in a non-conducting material situated inside an earthed conducting sheath BB. The insulating material was in some cases ebonite, and in others paraffin, but it was always found that the presence of the radium salt R caused a negative charge to accumulate progressively upon P.

Since the radium loses negative electricity on account of the β rays more readily than it does positive electricity carried away by the α rays, the β rays being the more penetrating and escaping more readily than the α rays, it would appear that radium enclosed in a non-conducting vessel would acquire a continually increasing positive charge. Lord Rayleigh, then Prof. Strutt,[1] constructed an interesting arrangement to exhibit this effect. The radium salt is contained in a tube A (Fig. 421) suspended in a vacuum tube, and therefore insulated from its surroundings. Attached to A is a pair of gold leaves. The walls of A are of such a thickness that the β rays can penetrate them and escape, but the α rays cannot. As A acquires a positive charge the gold leaves gradually diverge, until on touching the sides of the tube they are discharged, the process then starting afresh. Since the periodic time for the process is independent of external conditions, it is practically constant, and may be used to mark intervals of time. Such an apparatus would continue to act as long as the emission of β rays lasts, and this in the case of radium is measured in hundreds of years.

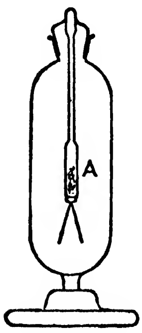


FIG. 421.

γ **Rays.**—These rays differ greatly from the two other kinds. They are non-deviable in a magnetic field, and do not carry an electric charge. Their chief characteristics are great penetrating power and ability to produce ionisation. Hence, the resemblance between the γ rays and the X-rays from a "hard" vacuum tube is very strong.

Of the fact that γ rays are of the same nature as X-rays there is no longer any doubt. Their mass absorption coefficients in aluminium (p. 486) have been measured in many cases by Rutherford and Richardson,[2] who found them to correspond to the series K and series L radiations for the metals of atomic weights equal to those from which the rays arise. The wave-length of the γ rays has been measured by Rutherford and Andrade [3] in the case of Ra B and Ra C, using a rock-salt crystal, and employing the X-ray spectrometer (p. 492). They also used a method

[1] R. J. Strutt, *Phil. Mag.*, **6**, p. 588. 1903.

[2] E. Rutherford and H. Richardson, *Phil. Mag.*, **25**, p. 722; **26**, pp. 324 and 937. 1913.

[3] E. Rutherford and E. N. da C. Andrade, *Phil. Mag.*, **27**, p. 854, and **28**, p. 263. 1914.

of transmission normally through a plate of the crystal, the rays finding their own appropriate reflecting planes. This gives concentration in certain directions in the transmitted beam, from which the angle of reflection could be calculated. The wave-lengths were found to lie between $1{\cdot}365 \times 10^{-8}$ and $7{\cdot}1 \times 10^{-9}$ cm.

In Fig. 422 is shown a diagram illustrating the deviability of α, β and γ rays in a magnetic field, first given by Mme. Curie. The γ rays are undeviated, while the deviation of the α rays to one side indicates their positive charge, the negative charge of the β rays being indicated by their deviation to the other side. The relative dispersions are also evident.

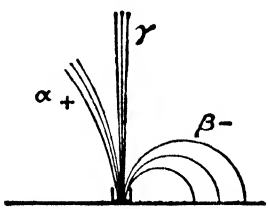


FIG. 422.

δ **Rays.**—In addition to the α, β and γ rays, slowly moving negative corpuscles have been detected by Sir J. J. Thomson [1] by means of the charge they carry. They would thus be similar in character to the β rays, but, owing to their lower velocity, they do not produce ionisation. Measurements of their velocity have shown this to be $3{\cdot}25 \times 10^8$ cm. per second, whereas the limiting velocity for the production of ionisation has been estimated to be $3{\cdot}6 \times 10^8$ cm. per second.

The Curie.—The standard of radioactivity is the quantity of radium emanation in equilibrium with 1 gramme of radium (p. 532), and is called the *curie*. It has a volume of 0·59 cubic mm. at standard temperature and pressure.

Radioactive Changes.—The emission of Becquerel rays by a radioactive substance is accompanied by a change or series of changes in the nature of the substance, changes both in its physical and its chemical properties, so profound and complex that their study has enormously increased our knowledge of the constitution of matter itself. The case of radium is typical. If a quantity of radium bromide be heated or dissolved in water, a new substance, gaseous in form, is separated from it, and immediately after the separation this new substance possesses very high radioactivity, while that of the radium is correspondingly reduced. If these two be examined after the lapse of a few days, it will be found that the activity of the radium has increased, while that of the other substance, known as its emanation, has fallen. The decay of activity of the emanation follows an exponential law, that is, the rate of decay is proportional to the activity; and at the same time the activity of the radium has increased according to a similar law. After a sufficiently long interval, the activity of the radium is completely restored, while

[1] J. J. Thomson, *Camb. Phil. Soc., Proc.*, **18**, p. 49. 1905.

that of the emanation, in fact the emanation itself, has entirely disappeared.

Uranium X.—Sir William Crookes [1] precipitated uranium from solution by means of ammonium carbonate, and redissolved the uranium by excess of the carbonate. A slight precipitate remained, which he found was several hundred times more photographically active than the original uranium. This substance he named uranium X. The photographic activity of Becquerel rays is chiefly due to the presence of β rays; hence the β rays are now emitted by the U X, and no longer by the uranium. Had the examination been by means of the power of producing ionisation, which is due to the α rays, it would have been found that the uranium still possesses this power, but not the U X.

It was subsequently shown by Becquerel that after the lapse of a year U X has completely lost its activity, while the uranium has regained its original condition.

Thorium X.—Radioactive processes are now known as those in which a chemically new substance is formed from some other, the change being generally accompanied with the emission of rays α, β or γ. The new substance generally decays at a rate represented by a logarithmic curve, and its production goes on at a constant rate within the parent material. Thus, on robbing the material of the new substance stored in it, its radioactivity is reduced at first by exactly the amount of that due to this stored material, but on being then allowed to remain undisturbed, its radiation will increase owing to the production of new material, until the loss by decay is balanced by the further production, in which case a condition of equilibrium is reached.

This explanation was put forward by Rutherford and Soddy [2] to account for the changes occurring in thorium, and it has subsequently been found that a similar explanation may be given to all radioactive changes, although the new substance formed may be solid, liquid or gas, its rate of decay may be rapid or slow, and the change may be accompanied by radiation or may be rayless. They precipitated the thorium from solution by means of ammonia, and found that the solution, which is free from thorium, has the greater part of the activity; and on evaporating to dryness and driving off the ammonium salts, a solid residue is obtained which in proportion to its weight is several thousand times as active as the original thorium. This substance was named thorium X, or Th X. After the lapse of a month the Th X had lost its activity, while the Th had completely recovered. On measuring the activity of the Th and Th X at known intervals

[1] W. Crookes, *Proc. Roy. Soc.*, **66**, p. 409. 1900.
[2] E. Rutherford and F. Soddy, *Phil. Mag.*, **4**, p. 370. 1902.

after their separation, the curves of Fig. 423 were obtained. It will be seen that, apart from slight irregularities at the start, the Th X loses its activity exponentially, that is, according to the law $\frac{A_T}{A_0}=\epsilon^{-\lambda T}$, where A_0 is the activity at the start, and A_T that after time T, λ being a constant. Moreover, the Th, which has only about 25 per cent. of its activity remaining on removal of

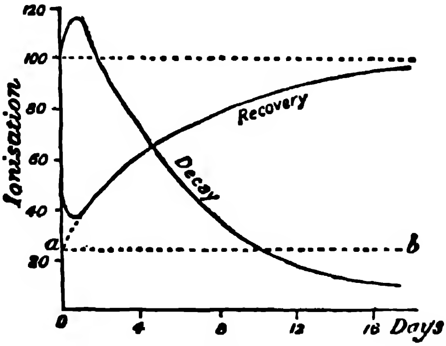


FIG. 423.

the Th X, regains its activity at a rate equal to the rate of loss of activity of the Th X. Hence, for the activity A_T recovered in time T,

$$\frac{A_T}{A_0}=1-\epsilon^{-\lambda T},$$

where A_0 is the activity recovered after infinite time. If the curve of recovery in Fig. 423 be measured from the dotted line *ab*, its equation will be found to fit approximately the curve. The activity of the Th X falls to half its value in about four days, and in the same time the Th performs half its recovery.

Let the whole mass of the thorium present produce a number q_0 of Th X atoms per unit time, and the rate of emission of activity by the Th X atom be K. Then, in order to find the total activity due to the Th X stored in the thorium after time T from separation, consider an interval of time dt after t seconds from the separation (Fig. 424). The number of Th X atoms produced in time dt is $q_0 dt$, and this has activity $Kq_0 dt$, which in the remaining interval $(T-t)$ decays in the ratio $\epsilon^{-\lambda(T-t)}$. Thus the activity at time T due to the Th X produced in the

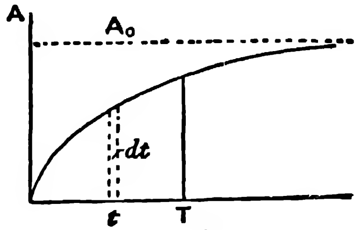


FIG. 424.

given interval dt is $Kq_0 . dt . \epsilon^{-\lambda(T-t)}$, and calling this dA we have—

$$dA = Kq_0 . \epsilon^{-\lambda(T-t)}dt,$$

and for the activity A_T of the whole of the Th X produced in the interval from 0 to T,

$$A_T = \int_0^T dA = \int_0^T Kq_0\epsilon^{-\lambda(T-t)}dt$$

$$= \frac{Kq_0}{\lambda}\Big[\epsilon^{-\lambda(T-t)}\Big]_0^T$$

$$= \frac{Kq_0}{\lambda}(1-\epsilon^{-\lambda T})$$

But when $T=\infty$, $A_T = A_0$,

$$\therefore A_0 = \frac{Kq_0}{\lambda},$$

or,

$$\frac{A_T}{A_0} = (1-\epsilon^{-\lambda T}),$$

an equation which is completely in accord with the experimental curve in Fig. 423.

Further, if the activity recovers by half the final amount in 4 days or 96 hours,

$$0{\cdot}5 = 1 - \epsilon^{-96\lambda},$$

from which, $\lambda = 0{\cdot}0072$,

or, if time be reckoned in seconds instead of hours—

$$\lambda = 2 \times 10^{-6}.$$

Similar measurements made upon uranium show that U X decays exponentially and U recovers in the same manner. In this case the time for half decay or recovery is 22 days, therefore

$$\lambda = 3{\cdot}6 \times 10^{-7}.$$

These constants are independent of the physical condition or state of chemical combination of the materials, for they are the same whatever the salt of uranium or thorium employed. The changes go on in exactly the same manner and at the same rate at the lowest and highest temperatures that can be employed. Hence they are changes occurring in the atom itself, and are independent of its motion and of its relation to other atoms.

Recovery when Parent Substance has Rapid Decay.—In the last case considered, the parent substance, thorium, decays so slowly that the Th X may be considered to be produced at constant rate; that is, q_0 is constant. There are, however, many cases in which the rate of decay of the parent substance is so rapid

that it must be taken into account. If thus q_0 is the number of atoms produced per unit time at the moment after the product has been all removed from the parent substance, this rate of production will fall to $q_0\epsilon^{-\lambda_1 t}$ in the interval of time t, owing to decay, where λ_1 is the radioactive constant of the parent substance. The number of atoms produced in time dt is then $q_0\epsilon^{-\lambda_1 t}dt$, and their activity is $Kq_0\epsilon^{-\lambda_1 t}dt$ as before. This will decay to $Kq_0\epsilon^{-\lambda_1 t}\epsilon^{-\lambda_2(T-t)}dt$ in the interval from t to T, where λ_2 is the radioactive constant of the substance produced,

then, $$dA=Kq_0\epsilon^{-\lambda_1 t}\epsilon^{-\lambda_2(T-t)}dt,$$

and the activity A_t of the whole of the product produced in the interval from 0 to T is—

$$A_t=\int_0^T dA=\int_0^T Kq_0\epsilon^{-\lambda_1 t}\epsilon^{-\lambda_2(T-t)}dt$$
$$=Kq_0\epsilon^{-\lambda_2 T}\int_0^T \epsilon^{-(\lambda_1-\lambda_2)t}dt$$
$$=\frac{Kq_0}{\lambda_2-\lambda_1}(\epsilon^{-\lambda_1 T}-\epsilon^{-\lambda_2 T})$$

It follows that $A_t=0$ when $T=0$, and again when $T=\infty$. The curve in Fig. 425 is calculated for Ac B.

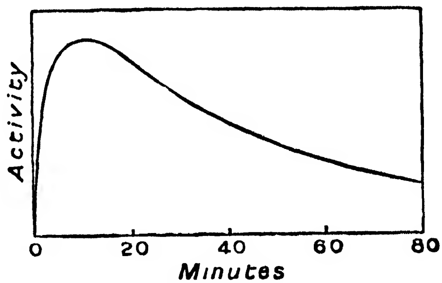


FIG. 425.

Radioactive Constant.—A definite meaning may be given to the constant λ, according to the above theory; it is the fraction of the amount of the product present which decays in unit time, and is called the *radioactive constant* of the product.

For the activity is measured by the ionisation produced, and this is almost entirely due to the a rays. Assuming that the ionisation produced by one a particle is constant, and that every atom as it changes projects the same number of a particles—

$$\frac{n_t}{n_0}=\frac{A_t}{A_0}=\epsilon^{-\lambda t},$$

where n_t and n_0 are the number per sec. of atoms changing respectively at time t, and when in radioactive equilibrium respectively.

Now, if N_t be the numbers of atoms of the product remaining after time t from separation from the parent substance—

$$n_t=\frac{dN_t}{dt}, \text{ or, } N_t=\int_t^\infty n_t dt,$$

since the number N_t will all subsequently change in the interval between t and ∞.

$$\therefore\ N_t=\int_t^{\infty} n_0\epsilon^{-\lambda t}dt=\frac{n_0}{\lambda}\epsilon^{-\lambda t}.$$

Hence the number N_0 at time $t=0$, or the number present for radioactive equilibrium, is—

$$N_0=\frac{n_0}{\lambda}.$$

Again, when equilibrium is reached, the number of atoms decaying per second is equal to the number produced per second, or—

$$n_0=q_0$$

so that, $$N_0=\frac{q_0}{\lambda}, \text{ or, } \lambda=\frac{q_0}{N_0},$$

and λ, the radioactive constant, is the ratio of the number of atoms changing per second to the number present.

Radioactive Constant and Range of α Rays.—An investigation of Geiger and Nuttall [1] brought to light an important relation between the average life of the atoms from which the α particle arises, and the velocity of emission and hence the range (p. 519) of the α particle. On plotting the logarithm of the radioactive constant (λ) for the various substances against the logarithms of the range (R) of the α particles emitted, a series of straight lines is obtained. The relation,

$$\log \lambda=A+B \log R,$$

is deduced, where A and B are constants. In cases where the substance is of too transient a nature for its radioactive constant to be determined by direct experiment, this may be found from the above relation when the range of its α rays has been found. In this way the value of decay to half-value for Th C′ and Ac A, given in the table on p. 538, were found.

Thorium Emanation.—From his experiments upon thorium, Rutherford [2] found that a substance in very minute quantity and having a gaseous form is given off by compounds of thorium, and may be carried into the vessel in which the ionisation is measured, by drawing the air from the neighbourhood of the thorium into the ionisation chamber. This effect is not due to the ionisation of the gas produced in the neighbourhood of the thorium and carried along by the moving air, since passing through porous material does not remove it (see p. 476), and if the thorium is wrapped in paper to absorb the α rays, which

[1] H. Geiger and J. M. Nuttall, *Phil. Mag.*, 22, 1912; 23 and 24, 1913.
[2] E. Rutherford, *Phil. Mag.*, 49, p. 1 1900.

produce most of the ionisation, the substance readily diffuses through the paper. The discoverer called this an *emanation*. Rutherford and Soddy [1] investigated the emanation as follows.

Air bubbled through strong sulphuric acid in A (Fig. 426), passes over the active material, wrapped in paper, in the tube B, through the bulb C, packed with cotton-wool, and into the ionisation vessel containing the three separate conductors D, E and F, either of which may be connected to the electrometer. The outer tube is joined to one pole of a battery, and the current from this to the electrometer measures the ionisation in the neighbourhood of D, E or F. The current is less the further the emanation has had to travel from the active material, so that for a given air velocity, the currents D, E and F form a diminishing series. By stopping the air current and measuring the fall of ionisation it was shown that the activity of the emanation falls exponentially, reaching half its value in 1 minute. Thus the radioactive constant in this case is given by $0{\cdot}5 = 1 \cdot \epsilon^{-60\lambda}$, or $\lambda = 1{\cdot}15 \times 10^{-2}$.

By placing a thorium salt, wrapped in paper to cut off direct

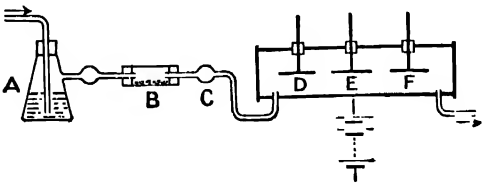


FIG. 426.

ionisation, in a closed vessel, the ionisation, and the current produced on account of it, rise owing to the production of the emanation. By measuring the ionisation current at intervals, it is found to increase exponentially, rising to half its value in one minute, the time of the decay of the emanation to one-half ; we may therefore conclude that the formation and decay of the emanation are related to each other in a similar manner to those of thorium X, the rate, however, being different.

It was shown by Rutherford and Soddy [2] that the thorium X, and not the thorium itself, is the source of the emanation, for, on removing the Th X by precipitating the thorium as on p. 524, the thorium has lost its power of producing emanation, while this is found now in the Th X. Moreover, the Th recovers its emanating power and the Th X loses it, both changes being half completed in 4 days. Thus the emanating power of Th X is proportional to its activity, and it is reasonable to conclude that

[1] E. Rutherford and F. Soddy, *Phil. Mag.*, **4**, p. 569. 1902.
[2] *Ibid.*

the Th X changes to emanation, the process being accompanied by the projection of α particles.

It should be remembered that Th X is a solid, and remains in the parent thorium, while the emanation is a gas and rapidly diffuses from the material. Rutherford and Soddy condensed the emanation in liquid air, in which case it adhered to the walls of the containing tube, passing off again in the gaseous form on rise of temperature. The condensation is not abrupt, beginning at about —120° C. and continuing over a range of about 30° to —150° C.

Excited Radioactivity.—It was found by Lord Rutherford [1] that a solid body placed in the emanation of thorium possesses a high radioactivity when removed from the emanation, and, further, that the amount of this excited activity is increased when the body is a negatively charged conductor, but the body does not receive any such activity if positively charged. The amount of excited activity is independent of the nature of the material upon which it is formed, except that when the body is to be negatively charged it is essential to use a conductor. That the excited activity is due to the emanation may be shown by covering the active thorium by a few sheets of paper to cut off the radiation of α particles, when the excited activity is still produced by the emanation which diffuses through the paper. If the thorium be covered by a sheet of mica sealed round the edges to keep in the emanation, the production of excited activity ceases.

That the excited activity is a product of change of the emanation may be shown by introducing a quantity of emanation into an ionisation vessel. The conductivity doubles in four or five hours, and if the emanation be then removed, the excited activity deposited on the walls of the vessel decays. Rutherford found that the excited activity produced by long exposure to the emanation decays to half in 11 hours, and that the recovery curve is related to it as the decay and recovery curves of thorium X.

The excited activity could be removed by dissolving in hydrochloric or sulphuric acids, but its mass is undetectable. It resides on the surface, since it may be partially removed by scraping.

Lord Rutherford gave the name of emanation X to the excited activity, since it is related to the emanation as thorium X is to thorium.

The thorium emanation itself emits only α rays, but after sufficient time has elapsed for appreciable formation of the excited activity, β and γ rays are also emitted. The emanation is enclosed in a copper vessel,[2] to absorb the α rays but to allow

[1] E. Rutherford, *Phil. Mag.*, **49**, p. 161. 1900.
[2] E. Rutherford and F. Soddy, *Phil. Mag.*, **5**, p. 445. 1903.

the β and γ rays to pass through. At first there is no ionisation outside the vessel, but after a time this does take place, and reaches a maximum in three or four hours. On removing the emanation by blowing it out of the vessel, this ionisation which is due to the β and γ rays emitted by the excited activity on the walls of the vessel does not disappear at once, but decays gradually.

The excited activity behaves differently according to whether the exposure of the body to the emanation is of short or long duration.[1] If short, the activity increases for several hours and then falls exponentially, but if the exposure has been long there is a fall from the start. Rutherford deduced from this behaviour that successive products were formed, which he called thorium A and thorium B. Subsequent investigations established that there is a whole succession of products, terminating in lead, which is stable and not radioactive. The sequence of changes can be followed from the table on p. 538 and the chart on p. 539.

At Th C there is a branching, some 65 per cent. of the atoms emitting a β particle to form Th C′, which in turn emits an α particle to form Th D, a form of lead, while the other 35 per cent. emit an α particle first, forming Th C″ which loses a β particle to form Th D. This branching phenomenon is found also in the other radioactive series.

Radioactive Equilibrium.—Since the number of atoms of any substance, of which the amount present is represented by X, transformed in one second is $\lambda_1 X$, where λ_1 is the radioactive constant (p. 527), the product, of which the amount present is Y, will not be in equilibrium with the parent substance until $\lambda_1 X = \lambda_2 Y$, when λ_2 is the radioactive constant of Y. For $\lambda_1 X$ being the number of atoms of X transformed per second, it is also the number of atoms of Y produced per second. Similarly for a series of products X, Y, Z, etc., in equilibrium—

$$\lambda_1 X = \lambda_2 Y = \lambda_3 Z, \text{ etc.}$$

This explains why a product of slow change such as uranium is present in minerals in much larger quantity than the more rapidly changing products.

Radium.—We have followed the radioactive changes taking place in the case of thorium in some detail, but it must be understood that the historical order of presentation has not been adhered to It will be simpler to trace the changes occurring in the case of other substances now that those for one have been described.

The discovery of radium has been described on p. 515.

The *emanation of radium* (*radon*) is not readily liberated from solid radium salts, but is occluded by them. On heating the salt,

[1] E. Rutherford, *Phil. Mag.*, **5**, p. 95. 1903.

the emanation is liberated, and a similar result is obtained by dissolving the salt in water.

Radium emanation is produced by a direct change from the radium itself, and not from a product of radium. Thus there is no radium product corresponding to thorium X. The emanation decays to one-half in 3·825 days. It is also condensed at low temperature,[1] the temperature of condensation being about −155° C.

The *excited activity* of radium emanation is similar to that met with in the case of thorium, and was first noticed by M. and Mme. Curie,[2] who showed that this is not due to the direct radiation from radium, but to the emanation, since it was produced on the solid when shielded from the direct radiation of the active salt. Like the excited activity of thorium, the deposit takes place much more readily upon a negatively electrified conductor than upon an uncharged body, but, unlike the case of thorium, a slight deposit will take place upon a positively charged conductor.

The decay curves for short exposure are complicated, and Rutherford concluded that, as in the case of thorium, there is a whole series of products. The sequence is given in the table of p. 538 and the chart of p. 539. It will be seen that radium itself arises from the decay of uranium, there being several intermediate substances, and that the subsequent decay sequence exhibits branching at Ra C not unlike that of Th C (p. 531), except that very few (about 0·04 per cent.) of the atoms decay *via* Ra C″.

It should be mentioned that the terminology for the various products has been changed as the story of the radioactive transformations was gradually unravelled and the letters used in the older papers, textbooks and books of tables do not always agree with the terminology used now.

Production of Radon.—The volume of radium emanation or *radon* from radium was measured by Ramsay and Soddy,[3] who collected the gases evolved in 8 days from a radium solution. These were largely oxygen and hydrogen, caused to combine by explosion, the excess hydrogen being removed by caustic soda. After drying by phosphorus pentoxide, the residue of emanation was measured. The radon shrinks in amount exponentially, to half volume in about 4 days, but if a large proportion of the α particles remain in the vessel, the resulting addition of helium gas (see p. 536) may increase the total volume. About 0·6 cub. mm. of radon (at 76 mm. of mercury pressure) is in equilibrium with 1 gm. of radium.

Radon is used in therapeutic work. It is pumped off from a vessel containing a radium source and sealed in small containers.

[1] E. Rutherford and F. Soddy, *Phil. Mag.*, **5**, p. 445. 1903.
[2] P. Curie and Mme. Curie, *Comptes Rendus*, **129**, p. 714. 1899.
[3] W. Ramsay and F. Soddy, *Proc. Roy. Soc.*, **73**, p. 346. 1904.

Emission of Heat by Radium.—The fact that radium compounds are permanently at a higher temperature than their surroundings, and therefore that radium is constantly emitting heat, was first pointed out by Curie and Laborde.[1] This difference of temperature is of the order of 2° C., but depends of course on the rate of escape of the emitted heat from the radium, as determined by the size of the specimen and the nature of its immediate surroundings. The rate of emission of heat was measured by the Bunsen ice calorimeter, and also by finding the rate at which heat should be supplied electrically to a similar and similarly situated mass of non-active material to maintain the same temperature. They found that radium emits heat at the rate of about 100 calories per hour per gramme of radium.

At a later date it was found by M. and Mme. Curie, in conjunction with Prof. Dewar, that the rate of emission of heat is still the same when the temperature is reduced to that of liquid oxygen, but Prof. Dewar thought that at the temperature of liquid hydrogen the rate is slightly increased.

The heat emitted during several stages of the radioactive changes occurring in radium was measured by Lord Rutherford and H. T. Barnes,[2] who placed the material contained in a small tube in turn into two flasks containing dried air, connected by means of a differential manometer. The difference in level in the manometer produced by transferring the radioactive material from one flask to the other is a measure of the rate of emission of heat, and was calibrated by finding the current in a fine piece of platinum wire of known resistance, that will produce the same effect.

The emanation was then removed from the radium, when the rate of emission fell rapidly to a minimum of 30 per cent., and then rose gradually to its original value in about a month.

The emanation was then tested, and it was found that its rate of emission was exactly complementary to that of the de-emanated radium ; in fact, the heat emission follows the same changes as the activity as measured by the ionisation produced by the α rays, and hence it is concluded that the heat emission is proportional to the activity measured by the α rays, and that the expulsion of each α particle corresponds to a constant production of heat.

For the change from radium to the emanation, the activity, as measured by the rays, is about 25 %, and the heat emission also about 25 %, of the total emission ; for the change from emanation to Ra A, the percentages are respectively 33 and 41, and for the remaining changes 42 and 34.

[1] P. Curie and A. Laborde, *Comptes Rendus*, **136**, p. 673. 1903.
[2] E. Rutherford and H. T. Barnes, *Phil. Mag.*, **7**, p. 202. 1904.

The total emission of heat by the emanation during its whole life may then be found on the assumption that the heat emission is proportional to the number of atoms breaking up per second; for,

$$n_t = n_0 \epsilon^{-\lambda t},$$

$$\therefore \text{ total number of atoms} = \int_0^\infty n_t dt = n_0 \int_0^\infty \epsilon^{-\lambda t} dt = \frac{n_0}{\lambda}.$$

$$\text{Hence, total emission of heat} = \frac{\text{rate of emission at start}}{\lambda}.$$

Now, from Rutherford and Barnes' experiment, the rate of emission of heat by the emanation alone from 1 gramme of radium is about 40 % of the total emission, that is, about 40 calories per hour, or $\frac{1}{90}$ calorie per second. Now, $\lambda = 2{\cdot}1 \times 10^{-6}$,

$$\therefore \text{ total heat emission} = \frac{10^6}{2{\cdot}1 \times 90} = 5300 \text{ calories.}$$

And similarly in its whole life radium emits a quantity of heat of the order of 10^{10} calories. When it is remembered that the union of hydrogen and oxygen to form 1 gramme of water causes an evolution of 3900 calories, the enormous store of energy within the atom in the case of the radioactive substances will be realised.

The heat production is due to the absorption of the energy of the radiations emitted and it comes ultimately from the atoms of the parent radioactive element. Release of energy from transmutation of elements must be taken into account in estimating the age of the earth and sun. On the theory of Helmholtz, that the energy of the sun is derived from shrinking under its own gravitational forces, Lord Kelvin estimated the age to be about 100 million years, but this period becomes enormously greater when account is taken of release of atomic energy (see p. 511).

It was shown by Lord Rayleigh that 270 tons of radium in the interior of the earth could account for the observed temperature gradient near the surface (about 1° C. for each 100 feet depth). Our somewhat limited knowledge of the actual distribution of radioactive minerals throughout the earth is not inconsistent with the idea that radioactivity plays a major part in maintaining this temperature gradient.

The numbers now accepted for the heat given by one gramme radium and its products per hour are, Ra alone 25·1, emanation 28·6, Ra A 30·3, Ra B and Ra C 50·5, making in all 134·7 calories.

Number of α Particles emitted by Radium.—Great importance attaches to a knowledge of the absolute number of α particles emitted by radioactive materials in a given time, and several methods have been adopted to determine this quantity. The most obvious method is to count the number of scintillations produced by a known amount of radium when the α particles

fall upon a screen of zinc sulphide, and this method was adopted by Regner.[1] An uncertainty in the interpretation of the result arises from the fact that we are not sure that every α particle produces a scintillation. The problem has been attacked by an entirely different method by Rutherford and Geiger.[2] They observed the increase of conductivity of a gas as each α particle is shot into it. The active material in the form of the active deposit of radium, situated upon the tip of a conical glass rod, A (Fig. 427), emits α rays in all directions. Those which fall upon the small aperture in the glass tube B, covered with a thin layer of mica, enter the measuring chamber C. Since each α particle in its travel produces about 43,000 ions, the increase in conductivity of the gas in C produced by one α particle may under favourable conditions become evident. The potential difference between the wire D and the outside tube C is adjusted until the saturation current is flowing from one to the other. The presence of a few extra ions will then, if the sparking stage is on the point of being reached, produce a large increase in the current, and it was found that a comparatively large throw of the electrometer needle occurred irregularly, but on the average something like four times a minute. The number of these impulses for a considerable period is counted, and so the number of α particles entering C per minute is known. Since the α rays are emitted from the active material uniformly in all directions, the total number emitted in a given time is to the number passing through B in the ratio of 4π to the solid angle subtended by the aperture in B at the point A. The ratio of the activity of the specimen to that of one gramme of radium was obtained during the experiment by means of the ionisation produced by the γ rays, and it was found as a result that the radium C in one gramme of radium emits $3{\cdot}4 \times 10^{10}$ α particles per second. Hence, counting the four α particles emitted by radium and its products in the course of their changes, one gramme of radium emits in all $13{\cdot}6 \times 10^{10}$ α particles per second.

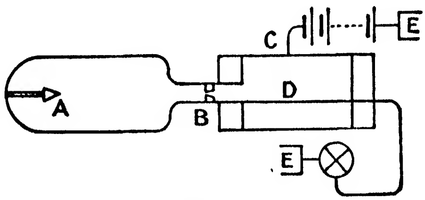


FIG. 427.

Charge on the α Particle.—Rutherford and Geiger also determined the total charge carried by the α particles from a deposit of radium C, by allowing this to fall on a conductor of known capacity, and finding the rise of potential per second, the β and δ rays being deflected away from the conductor by means of a

[1] Regner, *Sitz. Ber. der K. Preuss. Akad. der Wiss.*, **38**, 1909.
[2] E. Rutherford and H. Geiger, *Proc. Roy. Soc.*, A, **81**, pp. 141 and 162. 1908.

magnetic field. From their results, together with a knowledge from the previous experiment of the number of α particles emitted per second, it was found that each particle carries a charge of $9{\cdot}3\times10^{-10}$ electrostatic or $3{\cdot}1\times10^{-20}$ electromagnetic units, or twice the charge carried by a positive ion in the case of hydrogen in electrolysis. Now, $\frac{e'}{m'}$ is $4{\cdot}78\times10^{3}$ in electromagnetic units for the α rays (p. 518), and $\frac{e}{m}$ for hydrogen in electrolysis is $9{\cdot}6\times10^{3}$, and since e' is now seen to equal $2e$, it follows that $m'=\frac{9{\cdot}6}{4{\cdot}78}\cdot 2\cdot m=4{\cdot}02m$, and the atomic weight of the α particle is 4·02. That of the helium atom is 4·002, and hence it is probable that the helium atom is an α particle which has lost its positive or acquired a negative charge and become neutral.

It is now known quite definitely that the α particles are helium atoms with a charge of 2 positive electronic units. Also the amount of helium produced by 1 gramme of radium per year is 164 cubic millimetres.

Life of Radium.—Since the time that has lapsed since the discovery of radium is only an extremely small fraction of its whole life, it is impossible by direct observation of the rate of decay to determine this. Several computations by indirect methods have been made by Lord Rutherford.[1]

The rate of emission of α particles by a thin layer of radium bromide is found by measuring the current carried on their account to a neighbouring conductor, the experiment being performed at high vacuum to eliminate the disturbance due to ionisation of the gas, and with a strong transverse magnetic field to cause the slowly moving negative electrons to be bent back to their point of origin. The radium is at its minimum activity, the radioactive products having been removed, so that the β rays, which are only emitted by Ra B and Ra C, do not interfere. In this way it was found that if the charge on the α particle be taken as $1{\cdot}13\times10^{-20}$ electromagnetic units, that is, the charge on an electron, the total number of α particles emitted by 1 gramme of pure radium at its minimum activity is $6{\cdot}2\times10^{10}$. Using the active deposit upon lead, it was found in a similar manner that the Ra C in 1 gramme of radium emits $7{\cdot}3\times10^{10}$ β particles per second. This is probably slightly too high owing to the difficulty of ensuring that those which enter the supporting lead are absorbed, owing to their high power of penetration. Since one atom of radium in breaking down probably emits one α particle at the first stage and one β particle in all, it appears probable,

[1] E. Rutherford, *Phil. Mag.*, **10**, p. 193. 1905.

from these data, that in 1 gramme of radium $6{\cdot}2\times10^{10}$ atoms break up per second or $1{\cdot}95\times10^{18}$ per year. If 1 cubic centimetre of hydrogen contains $3{\cdot}6\times10^{19}$ molecules, and has a mass of $8{\cdot}96\times10^{-5}$ gramme, 1 gramme of hydrogen contains $\frac{2\times3{\cdot}6\times10^{19}}{8{\cdot}96\times10^{-5}}$ atoms, and the atomic weight of radium being 226, 1 gramme of it contains—

$$\frac{2\times3{\cdot}6\times10^{19}}{226\times8{\cdot}96\times10^{-5}}=3{\cdot}6\times10^{21} \text{ atoms,}$$

and the fraction breaking up in unit time or λ (p. 528) is—

$$\frac{1{\cdot}95\times10^{18}}{3{\cdot}6\times10^{21}}=5{\cdot}4\times10^{-4},$$

where the year is taken as the unit of time. This gives a decay to half in 1280 years, or an average life $\left(\frac{1}{\lambda}\right)$ of 1850 years.

Later experiments have shown that the charge on the α particle is twice that on the electron, and also that the number of α particles emitted by 1 gramme of radium per second in minimum activity is $3{\cdot}4\times10^{10}$, or $1{\cdot}07\times10^{18}$ per year, which would change the value of λ to $\frac{1{\cdot}07\times10^{18}}{3{\cdot}6\times10^{21}}=3\times10^{-4}$, and the decay to one-half occurs in 3300 years.

Another estimate is formed by considering the rate of emission of heat. From the mass and velocity of the α particle its kinetic energy would be about $5{\cdot}9\times10^{-6}$ ergs, and since 1 gramme of radium emits about 100 calories per hour,

$$5{\cdot}9\times10^{-6}\times n\times3600=100\times4{\cdot}2\times10^{+7},$$

or,

$$n=\frac{4{\cdot}2}{5{\cdot}9\times3{\cdot}6}\times10^{12}=2{\cdot}0\times10^{11},$$

on the assumption that the heat is produced by the bombardment of the α particles upon the substance and its surroundings, n being the total number emitted per second by radium and the radioactive products in equilibrium with it. For the radium in its condition of minimum activity the number is one quarter of this, that is, 5×10^{10} per second, which is in fair agreement with the number $3{\cdot}4\times10^{10}$ on p. 536 from the charge, and would, according to the above reasoning, mean a decay to half in 1600 years.

Further consideration of the volume of the emanation emitted in a given time (p. 532) led to a value of 1050 years for the period of decay to one-half. The value now accepted is 1590 years.

Table of Radioactive Substances.—The table of radioactive changes on p. 538 indicates the three great radioactive series of

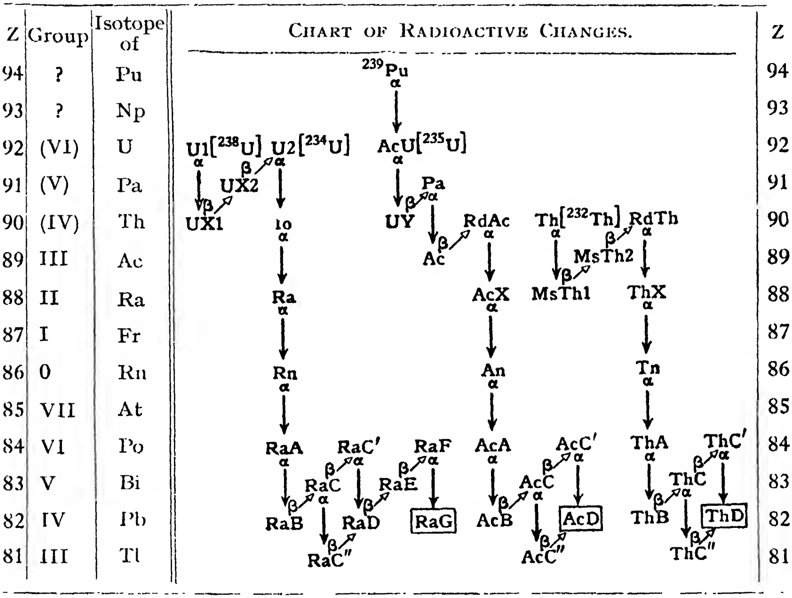
elements. There is little doubt that the final product of all the series is lead, although it is probable that the lead has not the same atomic weight in each case. Thus radium of atomic weight

TABLE OF RADIOACTIVE SUBSTANCES.

	Time for decay to half value.	Rays emitted.	Range of α particle in air at 76 cm. pressure and 15° C. in centimetres.
Uranium 1	$4{\cdot}6 \times 10^9$ years	α	2·68
Uranium X_1	24·5 days	β	—
Uranium X_2	1·14 min.	β	—
Uranium 2	3×10^5 years	α	3·24
Ionium	$8{\cdot}5 \times 10^4$ years	α	3·16
Radium	1590 years	α	3·26
Emanation (radon)	3·825 days	α	4·01
Radium A	3·0 min.	α	4·62
Radium B	26·8 min.	β	—
Radium C	19·7 min.	α, β	4·04
Radium C″	1·32 min.	β	—
Radium C′	10^{-6} sec. (?)	α	6·87
Radium D	22 years	β	—
Radium E	5 days	β	—
Radium F	139 days	α	3·81
Lead (Radium G)	—	—	—
Thorium	$1{\cdot}4 \times 10^{10}$ years	α	2·57
Mesothorium 1	6·7 years	β	—
Mesothorium 2	6·13 hours	β, γ	—
Radiothorium	1·9 years	α	3·90
Thorium X	3·64 days	α	4·24
Emanation (thoron)	54 sec.	α	4·97
Thorium A	0·14 sec.	α	5·60
Thorium B	10·5 hours	β	—
Thorium C	60·5 min.	α, β	4·68
Thorium C″	3·1 min.	β	—
Thorium C′	10^{-10} sec.	α	8·53
Lead (Thorium D)	—	—	—
Protoactinium	$3{\cdot}2 \times 10^4$ years	α	3·63
Actinium	13·5 years	—	—
Radioactinium	18·9 days	α	4·7
Actinium X	11·2 days	α	4·28
Emanation (actinon)	3·92 sec.	α	5·66
Actinium A	0·002 sec.	α	6·58
Actinium B	36·0 min.	β	—
Actinium C	2·16 min.	α, β	5·39
Actinium C′	10^{-3} sec.	α	6·52
Actinium C″	4·76 min.	β	—
Lead (Actinium D)	—	—	—

226 loses 5 α particles, or 5 helium atoms (atomic weight 4) in the course of the series of changes, and the end product should have the atomic weight 206. A determination of the atomic weight of lead from the uranium minerals gives it as 206·05. Similarly, thorium of atomic weight 232 loses in all 6 helium atoms, and we should expect the atomic weight of the end product to be 208. Experiment gives 207·9, and ordinary lead has atomic weight 207·2.

Isotopes.—The chemical and physical properties of an element depend largely on its atomic number Z. This quantity, which is discussed more fully in the next chapter, is the positive charge of the nucleus of the atom, expressed as a multiple of the electronic charge. Each electrically neutral atom has its nucleus surrounded by Z electrons and the arrangement of these electrons is a determining factor in chemical combination. As Z increases in unit steps, the properties of the element change systematically and it advances in the Periodic Table (see pp. 613 and 630). The loss of an α particle with its double charge lowers the atomic



number by 2 and places the element 2 places lower in the Table, as pointed out by Soddy,[1] while loss of a β particle corresponds to a gain of unity in Z (Fajans[2]). These changes are displayed in the accompanying chart, in which an obvious symbolism is used to denote loss of α and of β particles respectively.

The element shown as U1 starts the first radioactive sequence by loss first of one α particle, followed by 2 β particles, which restores its original nuclear charge. The first change, however, reduces its atomic mass by 4 while the β emissions have no appreciable effect. Hence U2 is 4 units lighter than U1 but has the same atomic number and hence the same number of electrons surrounding the nucleus and this confers on it the same properties. U1 and U2 are thus *isotopes* (cf. p. 503). All elements in the same

[1] F. Soddy, *Chem. News*, cvii, p. 97. 1913.
[2] K. Fajans, *Phys. Zeit.*, xiv, pp. 131 and 136. 1913.

horizontal row of the chart are isotopes. Those for $Z=86$ constitute the emanations and these are all inert gases placed in the Periodic Table in Group 0 with the other inert gases—helium, neon, argon, krypton and xenon. The final end-products of all three series are isotopes of lead, but the atomic masses are different —206, 207 and 208 respectively. These isotopes are all present in natural lead, which has an atomic weight 207·2.

Some other features of the Chart are discussed later.

Atmospheric Radioactivity.—In the period of 1900–1910 many observations were made which were interpreted in terms of ionisation by traces of radioactive substances in water, rocks, the air and various other substances. The discovery of cosmic rays, outlined on p. 507, showed that much of the observed radiation should in fact be attributed to these penetrating radiations coming from outside the earth. Nevertheless small amounts of natural radioactivity are quite widely spread. For example, a very small proportion of all the carbon in atmospheric carbon dioxide (of the order 1 in 10^{12}) consists of the weakly radioactive isotope of atomic weight 14, compared with 12 of normal carbon.

Nuclear Transmutation.—Radioactive changes, such as have been described in this chapter, involve a change in a nucleus which apparently occurs spontaneously, without any known external stimulus. In the common case of emission of an α particle, the change may be described as a partial disintegration. Other far-reaching changes may be produced by the impact on the nucleus of high-speed particles. Such particles are provided in nature in great quantity in cosmic rays, as already mentioned (p. 507).

Artificial transmutation of nuclei was first achieved by Rutherford,[1] who used α particles to bombard various light elements. Using α particles from Ra C, with a range in air of 7 cm., protons were ejected from hydrogen with ranges of up to 28 cm. and with even greater ranges from nitrogen. The protons were detected by scintillations and identified with the aid of their deflection in a magnetic field. It was shown in these experiments that the impinging particle attained a closest distance of approach estimated at $1{\cdot}9\times10^{-13}$ cm. when colliding with a hydrogen nucleus, and about double this for nitrogen. The protons derived from hydrogen may be regarded as being merely projected from the hydrogen molecules by direct impact, but with nitrogen the process involves nuclear change. In a nitrogen-filled Wilson Cloud Chamber, occasionally α particle tracks show a fork, the proton giving a long thin track and the recoiling nucleus a short thick one.

Protons of even longer ranges were later obtained by Rutherford and Chadwick[2] from other light elements. In the following

[1] E. Rutherford, *Phil. Mag.*, **37**, pp. 537, 562 and 581. 1919.
[2] E. Rutherford and J. Chadwick, *Phil. Mag.*, **42**, p. 809 (1921) and **44**, p. 417 (1922).

list, the atomic number is given with the symbol of the element, then the maximum range is given in cm.:

5, B—58; 7, N—40; 9, F—65; 11, Na—58; 12, Mg—40;
13, Al—90; 15, P—65.

By studying the relations between velocity and range for the particles concerned, Rutherford showed that, had the proton acquired the whole energy of the incident α particle, its range would be only about 57 cm. Hence in some at least of the examples quoted, part of the energy of the emitted proton must be derived from the transmutation of the nucleus.

Several other examples of nuclear transmutation will be discussed later.

The Neutron.—Bombardment by α particles proved to be a powerful weapon in studying atomic nuclei. In the course of such work, Bothe and Becker[1] detected a very penetrating radiation, particularly when beryllium was bombarded by α particles from polonium. These were at first thought to be very hard X-rays. Then Irène Curie and M. F. Joliot found[2] that although the rays penetrated metal sheets almost unaffected, the ionisation current which they produced in a suitable ionisation chamber could be almost doubled by the interposition of sheets of paraffin or other hydrogen-rich material. Thus the unknown rays projected hydrogen nuclei, *i.e.* protons, at high speed and so caused the extra ionisation.

Chadwick[3] established that the α-ray bombardment of the beryllium liberated a hitherto unknown particle which owed its long range to its absence of electric charge. This particle, the *neutron*, has a mass which is very close to that of the proton. Thus the impact of a neutron with a hydrogen nucleus is a problem of the collision of two practically equal bodies. To conserve both energy and momentum in such a case, if the impact is head-on, the particles virtually exchange velocities. Hence the neutron should be reduced to rest, while the proton carries on with the energy originally possessed by the neutron, thus explaining the appearance of high-speed protons from hydrogen-rich material.

The neutron produces no ionisation in a gas which it traverses and therefore leaves no visible track in a Wilson Cloud Chamber. A beam of neutrons can, however, be studied by the nuclear changes caused. Thus an ionisation chamber may be used filled with a suitable gas, boron fluoride, BF_3, being especially useful, or the chamber may have a lining of boron or lithium.

[1] W. Bothe and H. Becker, *Naturwiss.*, **18**, p. 705. 1930.
[2] I. Curie and M. F. Joliot, *Comptes Rendus*, cxciv, p. 273 and p. 708. 1932.
[3] J. Chadwick, *Nature*, **129**, p. 312. 1932.

The production of a neutron by α-bombardment of beryllium may be represented symbolically by the relation

$$ {}^{9}_{4}\mathrm{Be}+{}^{4}_{2}\mathrm{H} \rightarrow {}^{1}_{0}n+{}^{12}_{6}\mathrm{C}. $$

Here the superscript denotes the (approximate) mass of each particle in atomic mass units, while the subscript denotes the charge as a multiple of the electronic charge. Thus the neutron is shown as having unit mass and zero charge. The charges and masses on the two sides of the relation balance, although if accurate masses were used there would in general be some discrepancy in mass on account of the conversion of mass to energy which frequently occurs in nuclear reactions. As the residual nucleus in this example has charge $Z=6$, it must be an isotope of carbon.

The same reaction is sometimes represented by the more concise notation ${}^{9}_{4}\mathrm{Be}(\alpha, n){}^{12}_{6}\mathrm{C}$, in which the first particle represented inside the parentheses ( ) is the bombarding one, and the second is that emitted. The atomic number is sometimes omitted from these relations, since it is implied by the symbol for the element.

Induced or Artificial Radioactivity.—I. Curie and M. F. Joliot[1] found that, in certain cases, bombardment of light elements gave an emission of positrons (p. 509) and that this emission did not cease when the bombardment was suspended, but decayed according to the same exponential law as natural radioactivity. The half-value period for aluminium is $3\frac{1}{4}$ min., for magnesium $2\frac{1}{2}$ min., and for boron 14 min.

Presumably the impinging α particle enters the nucleus and a new nucleus is formed, with expulsion of a neutron:

$$ {}^{27}_{13}\mathrm{Al}+{}^{4}_{2}\mathrm{He} \rightarrow {}^{30}_{15}\mathrm{P}+{}^{1}_{0}n. $$

The new nucleus is an unstable isotope of phosphorus, unknown in nature, but it should have the same chemical behaviour as ordinary phosphorus. By carrying out appropriate reactions very quickly, Curie and Joliot showed that this was the case. For example, the gas generated from irradiated aluminium by a process which gives phosphine (PH_3) from ordinary phosphorus was found to emit positrons and the residue was inactive. Using β^+ to denote a positron, the decay may be represented

$$ {}^{30}_{15}\mathrm{P} \rightarrow {}^{30}_{14}\mathrm{Si}+{}^{0}_{1}\beta^+ $$

In a similar manner, with boron the reactions are:

$$ {}^{10}_{9}\mathrm{B}+{}^{4}_{2}\mathrm{He} \rightarrow {}^{13}_{7}\mathrm{N}+{}^{1}_{0}n; $$

$$ {}^{13}_{7}\mathrm{N} \rightarrow {}^{13}_{6}\mathrm{C}+{}^{0}_{1}\beta^+. $$

The final isotope in each case (${}^{30}\mathrm{Si}$, ${}^{13}\mathrm{C}$) is stable and known to occur in nature.

As will be seen later, the neutron plays a most important part in modern nuclear investigations.

[1] I. Curie and M. F. Joliot, *Comptes Rendus*, cxcviii, p. 254. 1934.

CHAPTER XVI

ATOMS

Faraday Tubes.—An ultimate explanation of physical phenomena is probably unattainable, but the order of development and abandonment of the successive hypotheses that have been employed to account for electrical and optical phenomena, and eventually for the two together, is an interesting and important study. The general tendency during the last century was to concentrate attention upon the medium in which matter is immersed, as the vehicle for the transference of energy from one place to another. Faraday explained electric and magnetic phenomena in terms of tubes of induction, and in optics the elastic solid theory, in spite of the great difficulties involved, was almost universally employed in explaining the propagation of the wave motion which we call light. The fundamental idea that the same medium may be used for the explanation of both electromagnetic and optical phenomena, in fact that light is an electromagnetic phenomenon, is due to Maxwell, but the further development, that the Faraday tubes with the properties attributed to them by Maxwell, alone are sufficient, is due to Sir J. J. Thomson, who developed this idea in his "Recent Researches in Electricity and Magnetism."

In Chapter XIII we found the velocity of propagation of an electromagnetic disturbance, from Maxwell's equations, to be $\frac{1}{\sqrt{k\mu}}$, and we afterwards used Thomson's Faraday tubes in the explanation of this propagation. If, however, a Faraday tube has a longitudinal tension, $\frac{2\pi D^2}{k}$ (p. 131), and be endowed with mass $4\pi\mu D^2$ per unit length (p. 432), then a lateral disturbance at any point would be propagated along the tube, as a lateral disturbance is propagated along a string under tension, the velocity of which is proved in works on Sound to be $\sqrt{\frac{F}{m}}$, where F is the tension, and m the mass per unit length. Now, in the case of a stretched string, the only force tending to restore the

string to its original shape is F the tension in it, there being no lateral pressure over its sides. In the case of the Faraday tube, however, we found on p. 134 that when the tube is curved there is a longitudinal restoring force $\frac{2\pi D^2}{k}$ due to its own tension, and also a component $\frac{2\pi D^2}{k}$ due to the lateral pressure of the neighbouring tubes, and thus F, in the problem of the string, must be replaced in the case of the Faraday tube by $\frac{4\pi D^2}{k}$. The velocity of propagation we should therefore expect to be

$$\sqrt{\frac{4\pi D^2}{k} \cdot \frac{1}{4\pi\mu D^2}} = \frac{1}{\sqrt{k\mu}},$$

which is entirely in accordance with the result obtained from Maxwell's equations.

Mass of Electrically-charged Sphere.—A uniformly charged sphere possesses a radial electrical field, the electric intensity E in its neighbourhood being $\frac{e}{kr^2}$, r being the distance of the point considered from the centre of the sphere, k the dielectric constant of the medium surrounding the sphere, and e the total charge. Thus, at every point, the electric displacement, or number of Faraday tubes per square centimetre, is $D = \frac{e}{4\pi r^2}$, the number of Faraday tubes arising upon the sphere being e. If the sphere be originally at rest we shall require some force to act upon it to put it in motion, quite apart from the question of its mechanical mass, for the Faraday tubes in motion possess energy, and on this account work must have been done. Let us consider the equivalent mass of the tubes in the small element of space at P (Fig. 428) due to the motion of the charged sphere in the direction Ox with velocity v. The angle between the direction of the tube and its velocity is θ, and therefore the equivalent mass per unit volume of the tube is $4\pi\mu D^2 \sin^2\theta$ (p. 432), that is $\frac{\mu e^2 \sin^2\theta}{4\pi r^4}$. Now, the area of the face of the element in the plane of the diagram is $r \,.\, d\theta \,.\, dr$, and for all such elements lying upon the circumference of the circle whose radius is QP

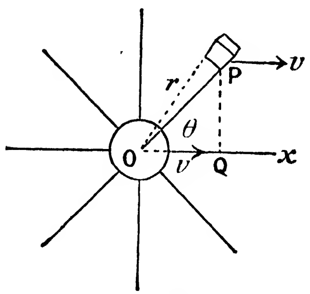


FIG. 428.

and whose plane is perpendicular to Ox, the electric intensity is the same, and consequently the mass of the tubes in the ring described by the area $rd\theta \,.\, dr$ if the diagram be rotated about Ox is—

$$\text{Volume of ring} \times \left(\frac{\mu e^2 \sin^2 \theta}{4\pi r^4}\right).$$

But,

$$\begin{aligned} \text{volume of ring} &= 2\pi \,.\, PQ \,.\, rd\theta \,.\, dr \\ &= 2\pi r \sin\theta \,.\, rd\theta \,.\, dr \\ &= 2\pi r^2 \sin\theta \,.\, d\theta \,.\, dr. \end{aligned}$$

Therefore, mass of ring due to the Faraday tubes

$$= 2\pi r^2 \sin\theta \,.\, d\theta \,.\, dr \,.\, \frac{\mu e^2 \sin^2\theta}{4\pi r^4}$$

$$= \frac{\mu e^2 \sin^3\theta \,.\, d\theta \,.\, dr}{2r^2}.$$

Hence for the whole of space surrounding the sphere,

$$\begin{aligned} \text{Mass due to Faraday tubes} &= \int_0^\pi \int_a^\infty \frac{\mu e^2 \sin^3\theta \,.\, d\theta dr}{2r^2} \\ &= \frac{\mu e^2}{2} \int_0^\pi \int_a^\infty \frac{\sin^3\theta}{r^2} \,.\, d\theta \,.\, dr, \end{aligned}$$

where a is the radius of the sphere.

Integrating first with respect to r, we have—

$$\int_a^\infty \frac{dr}{r^2} = -\left[\frac{1}{r}\right]_a^\infty = \frac{1}{a},$$

$$\therefore \text{mass} = \frac{\mu e^2}{2a} \int_0^\pi \sin^3\theta \,.\, d\theta.$$

$$\begin{aligned} \int_0^\pi \sin^3\theta \,.\, d\theta &= \int_0^\pi \sin\theta(1-\cos^2\theta)d\theta = -\int_{+1}^{-1}(1-\cos^2\theta)d\cos\theta \\ &= -\left[\cos\theta - \frac{\cos^3\theta}{3}\right]_{+1}^{-1} \\ &= -[-2+\tfrac{2}{3}] = \tfrac{4}{3}, \end{aligned}$$

$$\therefore \text{mass} = \frac{2\mu e^2}{3a}.$$

This mass must be added to any mechanical mass that the sphere may possess, in order to obtain its total mass. It has been calculated on the assumption that the Faraday tubes retain these symmetrical distributions when the sphere is put into motion. Now, this is not strictly true ; the Faraday tubes tend to set themselves at right angles to the direction of motion. Heaviside [1] has shown that for a charge moving with velocity v, the distribution of electric displacement is the same as though

[1] O. Heaviside, " Electromagnetic Theory," vol. 1.

the body were at rest, but the dielectric constant in the direction of motion was reduced in the ratio $\left(1-\frac{v^2}{c^2}\right)$, where c is the velocity of light. This effect would therefore be small for velocities much below that of light, and if the velocity of light could be attained, the value of k in this direction would be zero. Hence the Faraday tubes, as the velocity increases, tend to become more and more displaced towards the equatorial plane, that is, into a position in which their motion is at right angles to their length, and their effective mass therefore increases. In the limiting case when $v=c$ they are confined to an infinitely thin sheet at right angles to the direction of motion, the value of D, and likewise the electromagnetic mass of the charge, then being infinite.

Sir J. J. Thomson [1] has calculated the mass of a moving charge in terms of its velocity.

It is interesting to note that in finding the ratio $\frac{e}{m}$ for the β particles emitted by radioactive substances Kaufmann (p. 520) found the ratio to diminish as the velocity increased. Since it is very unlikely that e varies, we are driven to the conclusion that m increases, and Abraham,[2] on the assumption that the mass is entirely electromagnetic, calculated the deviations in electric and magnetic fields and found these to be very well in agreement with Kaufmann's observations.

More recently it has been shown to follow from the principles of relativity that the mass of any moving body is not independent of its velocity but increases in the ratio $\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ with increasing velocity v; whereas the linear dimensions in the direction of motion decrease in the ratio $\sqrt{1-\frac{v^2}{c^2}}$. Thus the dependence of mass upon velocity first deduced by Heaviside and Thomson from electromagnetic considerations in the case of moving charge is now known to be a special case of a universal principle applying to all matter.

Moving Charge equivalent to an Electric Current.—Since a charge moving with constant velocity is accompanied by its Faraday tubes moving with the same velocity, the magnetic field at any point may be expressed in terms of the movement of the tubes; in fact, the energy found for the moving tube on p. 432 is possessed on account of this magnetic field. On p. 422 we saw

[1] J. J. Thomson, " Recent Researches in Electricity and Magnetism."
[2] M. Abraham, *Phys. Zeitschr.*, 4, p. 57. 1902.

that the magnetic field is $4\pi Dv \sin\theta$. Now, in Fig. 428, the value of D at the point P is $\frac{e}{4\pi r^2}$, and therefore the magnetic field is

$$4\pi \cdot \frac{e}{4\pi r^2} \cdot v \sin\theta = \frac{ev \sin\theta}{r^2}.$$

It is directed from back to front if the moving charge e is positive and from front to back if e is negative. The value is the same for all points upon a circle having Q as centre and QP as radius, whose plane is at right angles to the direction of motion, and its direction is along the circle. The magnetic lines of force are therefore circles having their centres upon Ox and their planes at right angles to it.

On comparing this expression for the magnetic field with that due to a current element $i \cdot ds$ at point O, which, according to Ampère's rule given on p. 48 is $\frac{i \cdot ds \sin\theta}{r^2}$, we see that, for purposes of calculating magnetic field, the quantity ev for the moving electric charge is equivalent to $i \cdot ds$ for the continuous current.

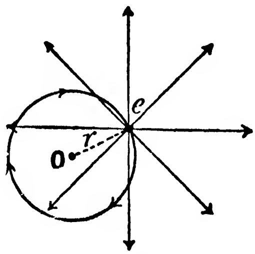


FIG. 429.

In the case of a charge e moving round a circle of radius r in periodic time T (Fig. 429), the Faraday tubes continually pass the centre of the circle O with velocity $\frac{2\pi r}{T}$, since each tube at the instant of passing O is moving at right angles to its own length and has the same velocity as e.

Now, at O the value of D is $\frac{e}{4\pi r^2}$,

$$\therefore \text{ magnetic field at } O = 4\pi Dv$$

$$= 4\pi \cdot \frac{e}{4\pi r^2} \cdot \frac{2\pi r}{T}$$

$$= \frac{2\pi e}{Tr}.$$

Now, a current i flowing in the circle has magnetic field $\frac{2\pi i}{r}$ at the centre, and hence for the moving charge to have the same magnetic field at O as the current—

$$\frac{e}{T} = i.$$

The magnetic moment of such an orbit is $\pi r^2 i = \frac{\pi r^2 e}{T}$.

Zeeman Effect.—It has been suggested that the motion of an electron in an orbit within an atom would give rise to pulsations in the tubes of electrical induction arising upon the electron. These pulsations would travel outwards with the velocity of light and might be considered to constitute the light radiation from the atom. The origin of the radiation is not in reality such a simple matter as this would suggest (see p. 595), but the consideration is of importance.

If the circular motion be resolved into two simple harmonic components parallel to OX and OZ (Fig. 430), the motion parallel to OX means a lateral displacement of the ends of those Faraday tubes which are parallel to OZ and OY, and hence disturbances travel in these directions. Similarly, the motion parallel to OZ causes waves to travel in the directions OX and OY. In the direction OY, the corresponding waves would at each point be represented by a rotating electric displacement, consisting of the two harmonic displacements, of which one is 90° in phase ahead of the other.

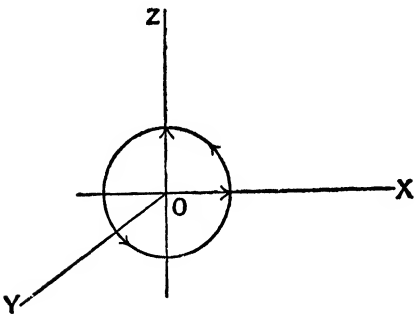


FIG. 430.

If the periodicity of the rotation of the charge is equal to that of light waves, the electromagnetic waves emitted are in all probability waves constituting light, and according to the electronic theory of Larmor and H. A. Lorentz light is due to the rotations of electrons within the atom. So long as these orbital motions of the electrons are undisturbed, it is extremely difficult to put the theory on an experimental basis, but the discovery of Zeeman [1] in 1896 that the light emitted by incandescent sodium vapour is modified by a magnetic field, made it seem probable. If a sodium burner be placed between the poles of a powerful electromagnet, and the emitted light analysed by means of a spectroscope, the D lines of sodium are both broadened, and if the resolving power be sufficient, are split up into several components. When the light is received in the direction of the magnetic field, having passed through a hole bored longitudinally through the pole pieces of the magnet, there are two components of the line, and these are found to be circularly polarised in opposite directions. When the light leaves the flame in a direction at right angles to the magnetic field there are three components; the middle one, which occupies the undisturbed position

[1] P. Zeeman, *Phil. Mag.*, **43**, p. 226. 1897.

of the line, is plane polarised, the direction of the electric displacement being parallel to that of the magnetic field, and the outer two are also plane polarised, but in a direction at right angles to the first.

These results are in accordance with the theory, and in the cases in which they are of the form described, the simple explanation given by Lorentz [1] suffices. In most cases, the behaviour of the spectral lines is more complicated, but the complete explanation has been given on the quantum theory.

Consider all the atoms concerned in the emission of light; their rotations may all be resolved into simple harmonic vibrations along the three rectangular axes OX, OY, and OZ (Fig. 431).

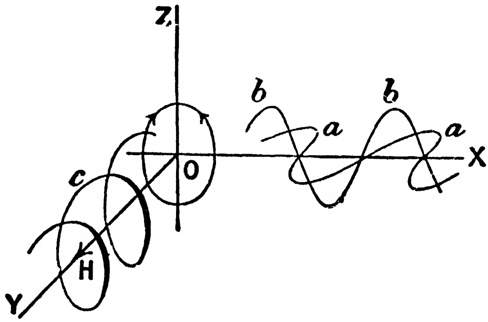


FIG. 431.

Those parallel to OY give rise to waves (a) spreading out in the plane ZOX, with direction of electric displacement parallel to OY, but not to waves travelling in the direction OY. The vibrations parallel to OZ and OX may, for convenience, be considered to be compounded into two equal and opposite rotational movements, which will give rise to plane polarised waves (b) in the plane ZOX, and to two circularly polarised waves, one of which is shown at c, travelling in the direction OY.

If now the magnetic field be in the direction OY, the vibration of the electric charge in this direction is unchanged, as are the waves (a). But the charges rotating in the circular paths are moving at right angles to the magnetic field, and consequently will experience forces of value Hev at right angles to their motion and to the field, the force on one being directed away from O, and on the other towards O. Assuming that the electron is subject to a restoring force acting towards the centre O of the atom and proportional to its distance r from O, there is equilibrium without the magnetic field when the centrifugal force $\frac{mv^2}{r}$

[1] H. A. Lorentz, *Rapports au Congrès Internationale de Physique*, 3, 1900.

is equal to the restoring force fr, f being the restoring force for unit displacement, and m the mass of the electron. If T be its period of vibration,

$$v=\frac{2\pi r}{T},$$

$$\therefore \frac{4\pi^2 rm}{T^2}=fr, \text{ or, } \frac{4\pi^2 m}{T^2}=f.$$

Now, the presence of the magnetic field produces a radial force Hev, and the restoring force fr must be increased or diminished by this amount according to the direction of rotation and the sign of the charge.

If then T_1 be the period of rotation in the magnetic field,

$$\frac{4\pi^2 r_1 m}{T_1^2}=fr_1 \pm Hev_1,$$

or,
$$\frac{4\pi^2 m}{T_1^2}=f\pm\frac{Hev_1}{r_1}, \text{ and since, } f=\frac{4\pi^2 m}{T^2},$$

$$\frac{4\pi^2 m}{T_1^2}-\frac{4\pi^2 m}{T^2}=\pm\frac{2\pi He}{T_1},$$

$$\therefore \frac{1}{T_1^2}-\frac{1}{T^2}=\pm\frac{He}{2\pi m T_1},$$

$$\frac{T^2-T_1^2}{T^4}=\pm\frac{He}{2\pi m T_1},$$

$$T-T_1=\pm\frac{HeT^2}{4\pi m} \quad \ldots\ldots\ldots \quad \text{(i)}$$

if T^2 be written for T_1T and $2T$ for $T+T_1$, both of which are justifiable when the changes in the periodic time produced by the magnetic field are small in comparison with T itself. Equation (i) may be written in terms of wave-lengths λ if we remember that $\lambda=cT$, and then becomes—

$$\lambda-\lambda_1=\pm\frac{He\lambda^2}{4\pi mc} \quad \ldots\ldots\ldots \quad \text{(ii)}$$

The periodic times of the two circular motions are therefore changed, one being diminished by the amount $\frac{HeT^2}{4\pi m}$, and the other increased by the same amount, and since each gives rise to a plane wave (b) we should expect there are two such waves emitted, one having greater frequency and the other less frequency than the undisturbed wave (a). Thus in the case of the light emitted at right angles to the magnetic field, the single

spectral line becomes a triplet, the central part being plane polarised with its electric displacement parallel to the magnetic field, and the two, one on either side of it, being plane polarised in a direction at right angles to this.

The light emitted in the direction of the magnetic field gives rise to a doublet, each component of which is circularly polarised. By means of a quarter-wave plate each of these circularly polarised rays can be rendered plane polarised, but the direction of the electric displacement depends upon the direction of rotation in the circularly polarised beam. With the direction of magnetic field given in Fig. 431, it is found that the anti-clockwise rotation, as seen on looking at the diagram, gives rise to the line displaced towards the blue end of the spectrum, and hence, applying the left-hand law (p. 240), we see that the rotating charge which gives rise to the light rays has the negative sign.

It is only when the vibration of the electron within the atom is of the simple harmonic type that we should expect the simple separation of the components of any spectral line described above. In nearly all cases the decomposition is into a much greater number of lines, but our explanation would show us that the light emitted at right angles to the field should always give plane polarised components, and the light along the direction of the field to oppositely circularly polarised components, and this is found to be the case. Moreover, whatever be the complexity of the resolved spectral line, the separation of the most widely separated components is a constant quantity.

According to Runge and Paschen,[1] for the normal triplet in the spectrum of mercury vapour, the quantity $\frac{\lambda_1-\lambda_2}{\lambda^2}$, where λ_1 and λ_2 are the wave-lengths of the displaced lines, was found to have the value 2·14 when the strength of magnetic field was 24,600. Hence, from equation (ii)—

$$\frac{\lambda_1-\lambda_2}{\lambda^2}=\frac{\mathrm{H}}{2\pi c}\cdot\frac{e}{m}=2{\cdot}14,$$

or,

$$\frac{e}{m}=\frac{2\pi\times 2{\cdot}14\times c}{\mathrm{H}}=\frac{6{\cdot}28\times 2{\cdot}14\times 3\times 10^{10}}{24{,}600}$$
$$=1{\cdot}6\times 10^{7}.$$

Now, the value of $\frac{e}{m}$ for the cathode rays is $1{\cdot}759\times 10^{7}$ (p. 483), and hence it is extremely likely that the constituent of the atom whose motion gives rise to the emission of light is the electron

[1] C Runge and F. Paschen, *Abhandl. der Berl. Akad.*, 1902.

met with in the cathode rays and in the β rays emitted by radioactive substances.

The Zeeman effect is generally more complicated than the above simple theory indicates, but it has been found that all the lines in similar series exhibit the same resolution in the magnetic field. The problem has also been treated by the quantum theory (pp. 595, 611), which has afforded a complete explanation of the effects observed.

Stark Effect.—It has been found by Stark[1] that an electric field can modify the light emission by atoms in a manner somewhat similar to the Zeeman effect. Positive rays are allowed to pass between two metal plates between which a strong electric field is maintained. On examining spectroscopically the light emitted he found that in the case of the Balmer series of hydrogen, each line is split into components whose number increases with the frequency, and the difference of frequency from that of the undisturbed line is proportional to the electric field. The light emitted at right angles to the field is plane polarised as in the Zeeman effect. The quantum theory has been applied successfully to the explanation of the Stark effect.

Dielectric Constant and Refractive Index.—The presence of atoms each consisting of some framework in which one or more electrons are situated, may be used to explain many of the physical properties of matter. The neutral atom has no resultant electric charge, but if it loses an electron, its resultant charge is positive. We know that the atom contains and is capable of losing at least one unit of electrical charge, the electron, which is equivalent to about $1{\cdot}59 \times 10^{-20}$ electromagnetic unit of charge. We may then draw a distinction between dielectrics or insulators on the one hand and conductors on the other; the former is a substance in which the electrons are displaceable within the atom, but are not detachable from it, while in the latter the electrons can be detached from the atoms, and are free to move in the spaces between them.

If f be the restoring force for unit displacement, and x the actual displacement of an electron from its neutral position, the restoring force is fx, and this is equal to the force Ee due to the electric field producing the displacement.

That is,
$$Ee = fx.$$

Within a mass of the material in which there are N atoms per unit volume, the component of the electric displacement due to one atom in each unit of volume being ex, that due to the N atoms is Nex. That corresponding to the original field is $\frac{E}{4\pi}$

[1] J. Stark, *Berliner Sitzungsber.*, Nov., 1913.

(p. 128), and hence the total electric displacement D within the material is the sum of the two quantities Nex and $\frac{E}{4\pi}$.

$$\therefore \; D=\frac{E}{4\pi}+Nex,$$

$$4\pi D=E+4\pi Nex$$
$$=E+\frac{4\pi NEe^2}{f}$$
$$=E\left(1+\frac{4\pi Ne^2}{f}\right).$$

Now, the dielectric constant $k=\frac{4\pi D}{E}$ (p. 128),

$$\therefore \; k=1+\frac{4\pi Ne^2}{f}.$$

Remembering that the refractive index $n=\frac{c}{v}$, where c is the velocity of light in vacuo and v is the velocity in the medium, and further that—

$$c=\frac{1}{\sqrt{k_0\mu_0}}, \text{ and, } v=\frac{1}{\sqrt{k\mu}},$$

we see that—

$$n=\sqrt{\frac{k\mu}{k_0\mu_0}}.$$

Now, for most substances μ is practically equal to μ_0, and if $k_0=1$ for vacuum—

$$n=\sqrt{\frac{k}{k_0}}, \text{ or, } n^2=\frac{k}{k_0}$$

and writing k for the ratio of the dielectric constant of the medium to k_0

$$n^2-1=\frac{4\pi Ne^2}{f}.$$

In the case of a gas, N is proportional to the density, and since the atoms are at considerable distances from each other, f is independent of the density,

$$\therefore \; n^2-1 \propto \text{density}.$$

On p. 429 we saw that for gases $k=n^2$. It was shown by Boltzmann[1] that the quantity $k-1$ is proportional to the pressure and therefore to the density of the gas. It should be

[1] L. Boltzmann, *Wien. Ber.*, **69**, 1874

noted that any theory which supposes the electric displacement within a dielectric to be due to the presence of small conducting bodies will lead to the above result.

Refraction and Dispersion.—Many theories have been advanced to account for the phenomena of the refraction and dispersion that occur as light passes from one medium to another. The mechanical theories, notably those of Fresnel and MacCulloch, afterwards modified by Sellmeier, account in a more or less satisfactory manner for the observed facts. But the electronic theory of H. A. Lorentz not only accounts for these same facts, but has the great advantage that it is in accordance with phenomena occurring in other branches of physics. As an example, we may cite the fact that the value of $\frac{e}{m}$ for the electrons concerned in the emission of light, as measured by the Zeeman phenomenon, is nearly identical with that found for the electrons in the cathode and the β rays. The mechanical theories suffered from the difficulty that it was almost impossible to reconcile the necessary properties of the incompressible æther with those of any known material substance.

On p. 439 we explained the reflection of electromagnetic waves from the surface of a conductor on the supposition that the electric intensity within the conductor is always zero. The reason for this last fact is clear, on the assumption that within the conductor free electrons exist which travel in a direction determined by that of the electric intensity.

If in Fig. 360, p. 439, the incident electrical intensity is directed upwards, and therefore on account of it the free electrons travel downwards, since their charges are negative, and the electric intensity due to them in their displaced position is directed downwards, that is, in opposition to the incident intensity; by their motion, as the harmonic wave arrives, they supply the equal and opposite harmonic variations of intensity which give rise to the two waves, one of which is the reflected wave, and the other a wave propagated into the conductor, and whose condition is everywhere equal and opposite to that of the incident wave, so that the two together have a zero resultant effect.

If the electrons are subject to some constraint, their motion will involve the expenditure of energy, and the reflection is not in this case perfect. For the best conductors known, the energy of the reflected beam is not quite equal to that of the incident beam. When the electrons are not free to move there will then be no reflection; the medium is perfectly transparent. No such substance is known; a certain amount of reflection always occurs as light passes from a vacuum to a material substance. But in the case of dielectrics, the electrons, although free to move within

the atom, are confined to the atoms, and it is only in the case in which their own proper periods of vibration within the atom approach that of the incident wave that any considerable amount of reflection occurs. The reflection in this case, although powerful, differs from the case of metallic reflection, in that only those waves having periodicity nearly equal to that of some particular free vibration of the electron within the atom are reflected.

Let us imagine that for a given substance the periodic time T_1 of the free vibrations of the electron within the atom is given by the equation $T_1=2\pi\sqrt{\frac{m}{f}}$, and let electromagnetic waves for which the electric intensity in the plane of incidence is $E=E_0 \sin 2\pi\frac{t}{T}$, fall upon the substance.

If x be the displacement of the electron from its position of equilibrium, the restoring force due to this is fx, and if the disturbing force due to the incident wave is Ee, the resultant force acting on it is $fx-Ee$, and the equation of motion of the electron is—

$$m\frac{d^2x}{dt^2}+fx-Ee=0,$$

or,

$$m\frac{d^2x}{dt^2}+fx=eE_0 \sin pt,$$

where,

$$\frac{2\pi}{T}=p,$$

$$\therefore \frac{d^2x}{dt^2}+\frac{f}{m}x=\frac{e}{m} . E_0 \sin pt.$$

Whatever its motion when the light is first incident upon it, it will after a few oscillations settle down to a steady vibration with the periodicity of the incident waves, and our problem is to find the amplitude of this vibration. The most general equation for this vibration is—

$$x=A \sin pt+B \cos pt,$$

where A and B are constants, at present unknown.

Hence,

$$\frac{dx}{dt}=Ap . \cos pt-Bp \sin pt,$$

$$\frac{d^2x}{dt^2}=-Ap^2 \sin pt-Bp^2 \cos pt,$$

and substituting the values for $\frac{d^2x}{dt^2}$ and x in the equation of motion, we have—

$$-Ap^2 \sin pt-Bp^2 \cos pt+\frac{f}{m}A \sin pt+\frac{f}{m}B \cos pt=\frac{e}{m}E_0 \sin pt.$$

This equation must be satisfied for all values of t. Now, when $pt=\frac{\pi}{2}$, $\cos pt=0$ and $\sin pt=1$,

$$\therefore\ -Ap^2+\frac{f}{m}A=\frac{e}{m}E_0,$$

and when $t=0$, $\sin pt=0$ and $\cos pt=1$.

$$\therefore\ -Bp^2+\frac{f}{m}B=0.$$

It follows that, except when $p^2=\frac{f}{m}$,

$$B=0,\ \text{and},\ A=\frac{\frac{e}{m}E_0}{\frac{f}{m}-p^2}.$$

Therefore the equation—

$$x=\frac{\frac{e}{m}E_0}{\frac{f}{m}-p^2}\sin pt$$

expresses the motion of the electron.

Now,
$$p=\frac{2\pi}{T},\ \text{and},\ \frac{f}{m}=\frac{4\pi^2}{T_1^2};$$

$$\therefore\ x=\frac{e}{m}\cdot\frac{E_0}{4\pi^2\left(\frac{1}{T_1^2}-\frac{1}{T^2}\right)}\sin pt.$$

The amplitude of vibration of the electron is therefore

$$\frac{e}{m}\cdot\frac{E_0}{4\pi^2\left(\frac{1}{T_1^2}-\frac{1}{T^2}\right)},\ \text{or},\ \frac{e}{m}\cdot\frac{E_0}{4\pi^2(n_1^2-n^2)},$$

where n_1 and n are corresponding frequencies.

We see, then, that when n_1 is very great, the amplitude of vibration due to the incident wave is infinitesimal, which means that if the electron is practically immovable, its presence does not affect the propagation of the wave. But if n_1 approaches in value to n, the amplitude increases and would, if our equation truly represented the facts, become infinite when $n_1=n$, which therefore corresponds to an ordinary case of resonance. We have, however, neglected all resistance to the motion of the electron, so that our result does not represent the truth when $n=n_1$.

The reasoning on p. 553 enables us to see the effect which, according to the electronic theory, this motion of the electrons will have upon the refractive index of the material.

Putting $D_0 \sin pt$ in place of D, $E_0 \sin pt$ in place of E, and

$$\frac{\frac{e}{m}E_0}{4\pi^2\left(\frac{1}{T_1^2}-\frac{1}{T^2}\right)}\sin pt \text{ in place of } x \text{ in the equation } 4\pi D = E + 4\pi Nex,$$

and dropping the term $\sin pt$ throughout, we have—

$$4\pi D_0 = E_0 + 4\pi\frac{Ne^2}{m} \cdot \frac{E_0}{4\pi^2\left(\frac{1}{T_1^2}-\frac{1}{T^2}\right)},$$

or,

$$k = 1 + \frac{Ne^2}{\pi m\left(\frac{1}{T_1^2}-\frac{1}{T^2}\right)} = n^2,$$

This may be expressed in terms of wave-lengths instead of periodic times, if by λ_1 we mean the wave-length in vacuo which corresponds to the periodic time T_1. Then $\lambda_1 = cT_1$ and $\lambda = cT$, so that—

$$n^2 = 1 + \frac{Ne^2\lambda_1^2}{\pi mc^2} \cdot \frac{\lambda^2}{\lambda^2 - \lambda_1^2} \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

where n is the index of refraction for the waves.

This equation will of course only represent the facts when the atoms are so far apart that the free period of the electron is not affected by the presence of neighbouring atoms, and when there is only one free period of oscillation. In the case of most substances there may be a great number of free periods, and each one will give rise to a term in the dispersion formula (equation iii).

The reason for the dispersion of light as it enters a refracting medium is now apparent, for the index of refraction depends upon the wave-length λ of the incident disturbance. When λ is very great in comparison with λ_1, the last equation becomes

$$n^2 = 1 + \frac{Ne^2\lambda_1^2}{\pi mc^2},$$

which is the refractive index for very long waves. Since $\lambda_1 = cT_1$ and $T_1 = 2\pi\sqrt{\frac{m}{f}}$, we see that this is the value found for n^2 on p. 553.

As λ approaches the value λ_1, the denominator of the last term in equation (iii) varies rapidly, and the value of n increases, becoming enormous when $\lambda = \lambda_1$. The curve AB in Fig. 432

indicates the manner in which n^2 varies as λ approaches λ_1, for a case in which n^2 is 1·25, when λ diminishes from infinity, and $\lambda_1=6\times10^{-5}$. When $\lambda<\lambda_1$ the sign of $\frac{\lambda^2}{\lambda^2-\lambda_1{}^2}$ changes, and the denominator increases while the numerator decreases. As therefore λ decreases, the value of n^2 increases from a large negative value, until when $\lambda=0$, $n=1$. This is indicated by the curve EF. Hence, for light waves of extremely short length the refractive index approaches the value unity.

In the case of a substance for which the electron has one free period of vibration, light waves of this period are strongly reflected, and hence the spectrum of the light transmitted by the substance exhibits an absorption band in this part of the spectrum; and, further, the light of this wave-length is that which will be emitted when the temperature is raised sufficiently for the body to be visible. This is a well-known result of Kirchhoff's laws of radiation.

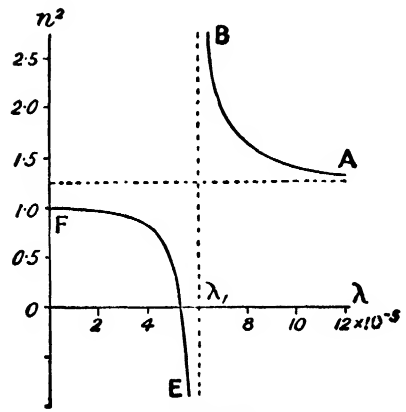


FIG. 432.

When the light absorbed has a much higher frequency than the visible rays, the refraction is "normal," as in the case of glass, etc., the waves of greater frequency being refracted most. The curve for n^2 and λ lies upon the part of AB (Fig. 432) remote from the origin. If, however, the absorption band occurs in the visible part of the spectrum, the two parts on either side of the band change places, the red part being refracted more than the blue, although in each part the natural order of the colours is preserved. The dispersion is then usually said to be "anomalous." Those substances, such as the aniline dyes, which have a strong absorption band in the visible part of the spectrum, exhibit the phenomenon of "anomalous" dispersion, and Professor R. W. Wood [1] has shown in a beautiful manner that this also occurs in the case of a prism of sodium vapour.

Faraday Effect.—Some connection between magnetism and light was suspected by Faraday, but the only connection which he was able to find was that the plane of polarisation of a beam of light passing through a dense transparent medium, such as a dense lead glass, experienced a rotation when the ray travels along the magnetic lines of force. On boring holes longitudinally

[1] R. W. Wood, *Roy. Soc. Proc.*, **69**, p. 157. 1901

through the poles of a powerful electromagnet, and passing a beam of light, plane polarised by means of a Nicol's prism, through these holes in such a way that it traverses a block of dense lead glass situated between the poles, it is found that the position of a second Nicol's prism, used as analyser to produce extinction of the beam, depends upon whether the magnetic field is on or off.

This phenomenon, known as the Faraday effect, bears some resemblance to the rotation of the plane of polarisation when a beam of light passes through certain crystallised media, such as quartz, in a direction parallel to the optic axis ; but there is one great difference between the two cases. In the crystal, the rotation of the plane of polarisation depends in some way upon the crystalline structure of the medium, and if on emergence the beam be reflected back so that it retraverses the crystal in the opposite direction, the opposite rotation occurs, and the beam emerges with its original direction of polarisation. But in the Faraday effect the direction of rotation is fixed in relation to that of the magnetic field, whatever be the direction of propagation of the beam, so that if the beam be caused to retrace its path through the field, the rotation of the plane of polarisation is doubled.

The rotation of the plane of polarisation depends upon the presence of the material medium in the magnetic field, and according to the electronic theory we can easily see why this should be. On p. 557 we saw that the index of refraction depends upon the free period of vibration of the electrons within the atom, and again on p. 550 we saw that the natural period of rotation of the electrons changes in the presence of a magnetic field, and also depends upon the direction of rotation.

Consider a plane polarised beam of light falling upon the block of dense glass ; in the absence of a magnetic field, the orbits of rotation of the electrons are circles, the two components of the motion at right angles to the direction of propagation being of equal frequency and clockwise and anti-clockwise respectively (see Fig. 431). We may, if we please, consider the plane vibration constituting the beam of light to be resolved into a clockwise and an anti-clockwise circular vibration, the relative phases of which determine the plane of polarisation (p. 376). These travel with equal velocities through the medium since the refractive index is modified by the electrons to the same extent for both components, and on emergence they are in the same relative phases as on entrance, and combine to form a plane polarised beam with its plane of vibration in the original direction.

If, however, there is a magnetic field, the electronic orbits are modified, one periodic time being increased and the other

diminished (p. 550), and the refractive indices for the two components of the beam are now different (p. 557). Hence they differ in phase on emergence by an amount depending upon the length of path in the medium and the strength of the magnetic field, and will combine into a plane polarised beam, whose plane of vibration is rotated from the original plane by an amount proportional to the difference of phase between the two circularly polarised components.

The rotation of the plane of polarisation was measured by Verdet for various substances and wave-lengths of light, and was found by him to fit the formula—

$$\theta = mlH\frac{n^2}{\lambda}\left(n - \lambda\frac{dn}{d\lambda}\right),$$

where l is the length of path traversed in the field, λ the wave-length in air of the light used, n the index of refraction, and m a constant depending upon the medium.

The quantity $\frac{\theta}{lH}$, or the rotation produced by travelling unit distance in unit field, is called Verdet's constant.

The following are a few of the values of Verdet's constant—

	Minutes of arc.	
Jena glass (densest silicate flint)	0·0888	(Du Bois, *Wied. Ann.*, **51**, 1893)
Quartz	0·01664	(Borel, *Arch. Genève*, (4), **16**, 1903)
Methyl alcohol	0·00989	(Quincke, *Wied. Ann.*, **24**, 1885)
Water	0·01311	(Roger and Watson, *Zeitsch. Phys. Chem.*, **19**, 1896)

Kerr Effect.[1]—It was found that plane polarised light, on being reflected from the polished pole of an electromagnet, was rendered elliptically polarised, and also that a transparent medium becomes slightly doubly refracting in a strong electrostatic field.[2]

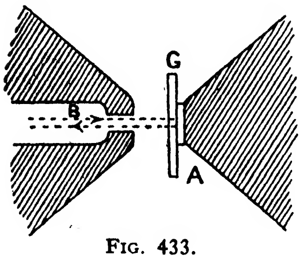


FIG. 433.

The Faraday and Kerr effects have been used by Du Bois [3] to measure strong magnetic fields and high intensities of magnetisation. A polished plate A (Fig. 433) of the metal under examination is placed on the flattened tip of the pole-piece of an electromagnet, and the plane polarised beam of light which

[1] J. Kerr, *Phil. Mag.*, **3**, p. 32 (1877); and **5**, p. 116 (1878).
[2] J. Kerr, *Phil. Mag.*, **50**, p. 337. 1875.
[3] H. E. J. G. du Bois, *Phil. Mag.*, **29**, p. 293. 1890.

traverses the hole B in the other pole of the magnet is reflected at A and returns along its own path. The rotation of the plane of polarisation is proportional to the intensity of magnetisation of A, the constant coefficient having been previously determined. The magnetising field is measured by placing a sheet of glass G, with its back surface silvered, to reflect the beam of light as before. From a knowledge of the rotation of the plane of polarisation and the value of Verdet's constant for the material, the strength of the magnetic field is known.

Paramagnetism and Diamagnetism.—The electronic theory has been adapted by Langevin [1] to account for the magnetic properties of materials. An electron rotating in a circular orbit is equivalent to a circular current, and corresponds very well to the hypothetical molecular currents of Ampere. Most atoms include many electrons, and the orbits of these may be so directed that the resultant magnetic moment for the atom is zero. In this case a magnetic field would not have any directive effect upon the atom. When, however, the resultant magnetic moment of the atom is not zero, either by reason of one electron uncompensated by another rotating in the opposite direction, or for any other reason, the atom possesses a resultant magnetic moment and consequently possesses energy in a magnetising field. If this energy is part of the thermal energy of the material the distribution of magnetic moments will be such that there is an increase in the magnetic induction (p. 567). That is, the permeability of the substance is greater than unity and its magnetic susceptibility is positive. Substances of this kind are said to be *paramagnetic*. The susceptibility is always comparatively small, as in the case of aluminium, platinum, etc. (p. 565). There is one small class of metals, iron, nickel and cobalt, in which the magnetic permeability is vastly greater than for any other substance. These are said to be the *ferromagnetic* metals. When the resultant magnetic moment of each atom is zero, the magnetising field will still have some effect, owing to its influence upon the electronic orbits themselves. The effect will be to decrease the magnetic induction, as shown below. Thus the magnetic permeability is less than unity, and the susceptibility is negative. Such substances are said to be *diamagnetic*, as, for example, copper, gold and bismuth (p. 565).

It is quite reasonable, therefore, to suppose that there are two opposite processes going on when a substance is placed in a magnetic field, the resulting susceptibility being a measure of the difference of the two. The orientation of the orbits (see p. 567) which will cause an increase in the magnetic induction is the first, and gives rise to paramagnetic properties. To understand

[1] P. Langevin, *Comptes Rendus*, **140**, p. 1171. 1905.

the second we may refer to a theory suggested by Weber and developed by Maxwell in his "Electricity and Magnetism." If the atom of the material be an electrical conductor, the establishment of a magnetic field will cause a current to flow in it. Let the circle AB (Fig. 434) be a line of flow of such a current within an atom, OH being the direction of the magnetic field. While the field is becoming established, this current is clockwise, as seen in the diagram, and hence the magnetic field within it, due to its own presence, is in the opposite direction to the original field. The magnetic induction is thus reduced by the presence of the conducting atom, and the presence of such conducting atoms would give to the material the diamagnetic property. If its electrical resistance be zero, the current, once started, will continue to flow until, on the removal of the magnetising field,

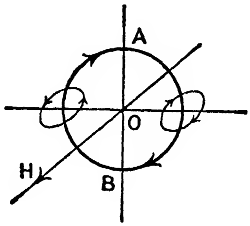


FIG. 434.

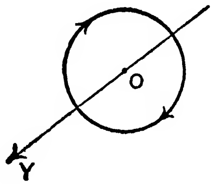


FIG. 435.

an opposite induced electromotive force reduces the current to zero.

If now we replace the conducting molecule of Weber and Maxwell by the atom with its rotating electrons, we can explain the para- and diamagnetism. For the rotation into the direction of the field increases the magnetic induction within the material; and we will now proceed to show that the alteration in the orbits produced by the magnetising field reduces the induction. The material will then be para- or diamagnetic according to which effect is the greater.

For a mass of the material, the possible motions of the electrons may all be resolved, as described on p. 549, into a linear vibration along one axis OY (Fig. 431) and two equal and opposite circular rotations in a plane at right angles to this. Let the magnetising field H be in the direction OY; then the linear vibration along this axis is unaffected by the field, but the circular motions are affected as already described. For the rotation of the electron shown by the arrow in Fig. 435, the equivalent current is in the opposite direction, since the electron is a negative charge; the force due to the field is Hev and is directed outwards along the radius, and hence, as on p. 550—

$$\frac{mv_1^2}{r_1}=fr_1-Hev_1,$$

or,
$$\frac{4\pi^2}{T_1^2}-\frac{f}{m}=-\frac{He}{m}\cdot\frac{2\pi}{T_1} \quad \ldots \ldots \quad (a)$$

Now, for the rotation of the electron in the opposite direction, the force due to the magnetic field is directed inwards, that is towards O,

$$\therefore \frac{mv_2^2}{r_2}=fr_2+Hev_2,$$

or,
$$\frac{4\pi^2}{T_2^2}-\frac{f}{m}=+\frac{He}{m}\cdot\frac{2\pi}{T_2} \quad \ldots \ldots \quad (b)$$

Subtracting equation (*b*) from (*a*), we have—

$$\frac{1}{T_1^2}-\frac{1}{T_2^2}=-\frac{He}{2\pi m}\left(\frac{1}{T_1}+\frac{1}{T_2}\right),$$

$$\therefore \frac{1}{T_1}-\frac{1}{T_2}=-\frac{He}{2\pi m} \quad \ldots \ldots \ldots \quad (c)$$

Again, we saw on p. 547 that a charge e moving in a circular orbit in periodic time T is equivalent to a circular current $\frac{e}{T}$, and hence the magnetic moment of such an orbit is $\frac{ea}{T}$, where a is the area of the circle. If there are N_1 such orbits per unit volume of the material, $\frac{N_1ea}{T}$ is the magnetic moment, or the intensity of magnetisation due to them. When the magnetising field is zero, the magnetic moment due to the electrons rotating in either direction is $\frac{N_1ea}{T}$, and these being oppositely directed, the resulting intensity of magnetisation is zero. In the presence of the magnetic field, the magnetic moment due to the electrons rotating as shown in Fig. 435 is $\frac{N_1ea}{T_1}$, and is directed from O to Y; that due to the opposite rotations is $\frac{N_1ea}{T_2}$, and is directed from Y to O.

From equation (*c*)—

$$\frac{N_1ea}{T_1}-\frac{N_1ea}{T_2}=-\frac{N_1He^2a}{2\pi m},$$

the left-hand side of which equation is the resultant intensity of magnetisation directed from O to Y, and the right-hand side

shows us that this is always negative, and hence that the magnetisation is in the opposite direction to the magnetising field. Also the ratio of intensity of magnetisation to magnetising field is the magnetic susceptibility κ;

$$\therefore \kappa = -\frac{N_1 e^2 a}{2\pi m}.$$

In this case N_1 is not the total number of rotating electrons in unit volume of the material. If we imagine all the rotating electrons to be divided into six groups having for their axis of rotation the three rectangular axes respectively, the two rotations about any axis being opposite, only those having axes parallel to the magnetic field have any magnetic moment in this direction, and therefore only two out of the six groups, that is one-third of the total number of rotating electrons present, are concerned in determining the susceptibility. Hence, if N is the total number present,

$$N_1 = \frac{N}{3},$$

and,

$$\kappa = -\frac{Ne^2 a}{6\pi m}.$$

We see, then, that the property of diamagnetism is the result of effects occurring within the atom, and would therefore be independent of temperature. Such is found to be the case with most substances, antimony and bismuth being exceptions. If the opposite rotations of the electrons within the atom are symmetrical, the diamagnetic property only would be exhibited, but if unsymmetrical there would be a resultant magnetic moment with exhibition of paramagnetism.

It was found by Curie [1] that the magnetic susceptibility referred to unit mass varies inversely as the absolute temperature in the case of the paramagnetic substances, and this is now known as Curie's law. The quantity χT, where $\chi = \frac{\kappa}{\text{density}}$ and T the absolute temperature, is Curie's constant. In the case of oxygen, Curie found the magnetic susceptibility to be represented very well by the relation $10^6 \chi = \frac{33{,}700}{T}$, and by measuring the force on small metallic spheres, first in air and then in liquid oxygen, Dewar and Fleming [2] found the value of κ for liquid oxygen to be 324×10^{-6}, or $\mu = 1{\cdot}00407$.

In the following tables are some of the magnetic constants for

[1] P. Curie, *Journ. d. Phys.*, **4**, p. 197. 1895.
[2] J. Dewar and J. A. Fleming, *Proc. Roy. Soc.*, **63**, p. 311. 1898.

the elements and common substances (see Kaye and Laby's Tables) :—

Aluminium	$\chi=+0{\cdot}65\ \times10^{-6}$
Bismuth	$-1{\cdot}38\ \times10^{-6}$
Copper	$-0{\cdot}09\ \times10^{-6}$
Gold	$-0{\cdot}15\ \times10^{-6}$
Platinum	$+1{\cdot}10\ \times10^{-6}$
Potassium	$+0{\cdot}52\ \times10^{-6}$
Silver	$-0{\cdot}2\ \ \times10^{-6}$
Mercury	$-0{\cdot}19\ \times10^{-6}$
Water	$-0{\cdot}72\ \times10^{-6}$
Air	$\kappa=+0{\cdot}029\times10^{-6}$
Hydrogen	$-0{\cdot}002\times10^{-7}$
Oxygen	$+0{\cdot}139\times10^{-6}$

In order to explain paramagnetism consider the atom to have electron orbits (or anything equivalent to a current circuit) which are not balanced. When situated in a magnetising field the orbits, which possess magnetic moment, will turn into the direction of the field. This, however, is not a simple turning such as would occur with a bar magnet, for the effect of a couple on a rotating body, such as an electron, would be to produce precession, as in the case of a spinning top. If, however, the energy of the atom in the magnet field be compared with the energy of thermal agitation in the kinetic theory of gases, an expression for the distribution of the atomic magnetic moments in space may be found.

In the case of a gas, in which the molecules are supposed to approximate to points, there are only three degrees of freedom, which correspond to motion parallel to three axes in space. If the gas as a whole is at rest, the resolved velocities parallel to the three rectangular axes, integrated for the molecules in unit volume are equal. The pressure of the gas is then $\frac{1}{3}\mathrm{N}mv^2$, where N is the number of molecules per c.c., m the mass of each molecule and v^2 the mean of the squares of the velocities of all the molecules, that is, $p=\frac{1}{3}\mathrm{N}mv^2$ (p. 576). The factor $\frac{1}{3}$ arises on account of the kinetic energy ($\frac{1}{2}\mathrm{N}mv^2$) being equally distributed amongst the velocities parallel to the three axes.

When the gravitational field in which the gas exists is taken into account, the molecules in a vertical column will no longer be uniformly distributed. In the lower levels the density of the gas is greater than in the upper levels. Let dp be the difference in pressure between two horizontal planes whose vertical distance apart is dx.

$$dp=\rho g dx$$

The density ρ is Nm, $\therefore dp=Nmgdx$. Consider volume V of the gas,

$$pV=RT=\tfrac{1}{3}Nmv^2V \qquad \text{(p. 576)}$$

and if V is 1 c.c., $\frac{1}{3}mv^2=\frac{RT}{N}$.

Again, since $p=\frac{1}{3}Nmv^2$, $dp=\frac{1}{3}mv^2\,.\,dN$ because, at constant temperature $\frac{1}{3}mv^2$ is constant, since $\frac{1}{2}mv^2$, or the mean kinetic energy of the molecules is constant. Although the kinetic energy may not be the same for all the molecules, the mean is maintained by perpetual collisions. From the two expressions for dp—

$$\tfrac{1}{3}mv^2\,.\,dN=\rho g dx$$

$$\frac{RT}{N}dN=\rho g dx$$

$$\frac{dN}{N}=\frac{\rho g}{RT}dx$$

On integrating, $$\Big[\log N\Big]_1^2=\frac{\rho g}{RT}(x_2-x_1)$$

$$\log\frac{N_2}{N_1}=\frac{\rho g(x_2-x_1)}{RT}$$

$\rho g(x_2-x_1)$ is the work done when 1 c.c. of gas is transferred from level x_1 to level x_2, or the difference in the potential energy of 1 c.c. of gas at the two levels. Calling this W, then $\log\frac{N_2}{N_1}=-\frac{W}{RT}$, or $\frac{N_2}{N_1}=\epsilon^{-\frac{W}{RT}}$. The negative sign is taken because N is greatest at the lower level, where the potential energy is least. It is common practice to express the difference in potential energy as the work required to transfer one molecule of the gas, instead of 1 c.c. from one level to the other, so that $w=\frac{W}{N}$. In this case R, the gas constant referring to 1 c.c., must be replaced by $\frac{R}{N}=k$, and

$$\frac{N_2}{N_1}=\epsilon^{-\frac{w}{kT}}$$

k is frequently known as the Boltzmann constant.

Turning now to the case of the distribution of atoms or parts of atoms of magnetic moment m, it is seen that the couple acting on each, when situated in field H, is $Hm\sin\theta$, where θ is the angle between the directions of the field and magnetic moment. For a rotation $d\theta$ the work done is $Hm\sin\theta\,.\,d\theta$. Taking the

potential energy of the atom as zero when its magnetic moment is at right angles to the field, so that its component of magnetic moment in the direction of the field is zero, the work done in turning from this position to one making angle θ with the field is $\int_{\frac{\pi}{2}}^{\theta} m\mathrm{H}\sin\theta \,.\, d\theta = -m\mathrm{H}\Big[\cos\theta\Big]_{\frac{\pi}{2}}^{\theta} = -m\mathrm{H}\cos\theta$. The potential energy varies from $-m\mathrm{H}$ when $\theta=0$ to $+m\mathrm{H}$ when $\theta=\pi$; it therefore increases as θ increases, just as, in the case of a gas the potential energy of the molecule increases as x increases. As a system tends to its condition of least potential energy, the tendency will be for θ to decrease in the presence of a magnetising field, that is, the material will exhibit paramagnetic properties.

For the N atoms in 1 c.c. whose moments make angle θ with the field, the potential energy is $-\mathrm{N}m\mathrm{H}\cos\theta = b\mathrm{RT}$, where b is a constant which depends on the fraction of the total temperature energy which is possessed by the body on account of this energy of magnetisation in the magnetising field. If N changes, then, change of energy per c.c. $= -m\cos\theta\mathrm{H} \,.\, d\mathrm{N} = d\mathrm{W}$

$$\therefore \frac{b\mathrm{RT}}{\mathrm{N}} d\mathrm{N} = d\mathrm{W}$$

On integrating,

$$\Big[\log \mathrm{N}\Big]_0^1 = \frac{\mathrm{W}}{b\mathrm{RT}}$$

$$\frac{\mathrm{N}_1}{\mathrm{N}_0} = \epsilon^{\frac{\mathrm{W}}{b\mathrm{RT}}}$$

or in terms of the energy for one atom instead of 1 c.c.

$$\frac{\mathrm{N}_1}{\mathrm{N}_0} = \epsilon^{\frac{w}{k\mathrm{T}}}$$

where $w = \frac{\mathrm{W}}{n}$ and $k = \frac{b\mathrm{R}}{n}$, n being the number of atoms per c.c. concerned in this part of the temperature energy.

Thus $\mathrm{N}_1 = \mathrm{N}_0\epsilon^{\frac{m\mathrm{H}\cos\theta}{k\mathrm{T}}} = \mathrm{N}_0\epsilon^{a\cos\theta}$, where $a = \frac{m\mathrm{H}}{k\mathrm{T}}$ and N_0 is the number of atoms with the arbitrary zero of potential energy. Now consider a sphere of radius r within the material, the centre being at O (Fig. 436), the number of atoms per unit volume situated at O, whose moment is directed along OP is $\mathrm{N}_0\epsilon^{a\cos\theta}$ and this, multiplied by the solid angle subtended by the strip QP will be the total number of atoms whose moments are directed towards the strip, that is,

$$\mathrm{N}_0\epsilon^{a\cos\theta} \times \frac{2\pi r^2 \sin\theta d\theta}{r^2} \quad \text{or} \quad 2\pi\mathrm{N}_0\epsilon^{a\cos\theta}\sin\theta \,.\, d\theta.$$

Integrating this from 0 to π gives the whole number of atoms in the unit volume. Calling this N, we have—

$$N=2\pi N_0\int_0^\pi \epsilon^{a\cos\theta}\sin\theta\,.\,d\theta$$

$$=-\frac{2\pi N_0}{a}\int_a^{-a}\epsilon^{(a\cos\theta)}d(a\cos\theta)$$

$$=-\frac{2\pi N_0}{a}(\epsilon^{-a}-\epsilon^{+a})$$

$$=\frac{2\pi N_0}{a}(\epsilon^{a}-\epsilon^{-a})=\frac{4\pi N_0}{a}\sinh a.$$

Again, the magnetic moment of one of these atoms being m, its component along OY is $m\cos\theta$, and the resultant magnetic

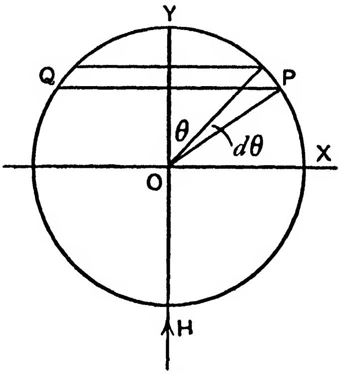


FIG. 436.

moment along OY of all the atoms whose moments are directed towards the strip QP is $2\pi N_0\epsilon^{a\cos\theta}\sin\theta\,.\,d\theta\times(m\cos\theta)$

$$=2\pi N_0 m\epsilon^{a\cos\theta}\sin\theta\,.\,\cos\theta\,.\,d\theta$$

and the resultant moment along OY for all the N atoms in the unit volume is

$$I=2\pi N_0 m\int_0^\pi \epsilon^{a\cos\theta}\sin\theta\,.\,\cos\theta\,.\,d\theta$$

$$=-\frac{2\pi N_0 m}{a^2}\int_0^\pi \epsilon^{a\cos\theta}a\cos\theta\,.\,d(a\cos\theta).$$

Integrating $\int\epsilon^x x dx$ by parts,

$$\int\epsilon^x x dx=\epsilon^x x-\int\epsilon^x dx=\epsilon^x x-\epsilon^x$$

$$\therefore\ I=-\frac{2\pi N_0 m}{a^2}\Big[\epsilon^{a\cos\theta}a\cos\theta-\epsilon^{a\cos\theta}\Big]_0^\pi$$

$$=\frac{2\pi N_0 m}{a^2}(a\epsilon^a+a\epsilon^{-a}-\epsilon^a+\epsilon^{-a})$$

$$=\frac{4\pi N_0 m}{a}\left(\frac{\epsilon^a+\epsilon^{-a}}{2}-\frac{\epsilon^a-\epsilon^{-a}}{2a}\right)$$

$$=\frac{4\pi N_0 m}{a}\left(\cosh a-\frac{1}{a}\sinh a\right).$$

If all the N atoms per c.c. of magnetic moment m were oriented in the direction of the magnetic field the intensity of magnetisation would be Nm, the saturation value. From the above values of I and N,

$$\frac{I}{Nm}=\frac{\cosh a}{\sinh a}-\frac{1}{a}$$

$$=\coth a-\frac{1}{a}$$

$$=\frac{\epsilon^a+\epsilon^{-a}}{\epsilon^a-\epsilon^{-a}}-\frac{1}{a},$$

The quantity $\frac{I}{Nm}$ takes the value $\frac{a}{3}$ when a is small, which may be proved by expanding ϵ^a in powers of a and neglecting a^4 and higher powers.

Thus,
$$\epsilon^a=1+a+\frac{a^2}{\underline{|2}}+\frac{a^3}{\underline{|3}}+\ldots$$

$$\epsilon^{-a}=1-a+\frac{a^2}{\underline{|2}}-\frac{a^3}{\underline{|3}}+\ldots$$

$$\therefore\quad \epsilon^a+\epsilon^{-a}=2\left(1+\frac{a^2}{\underline{|2}}+\frac{a^4}{\underline{|4}}+\ldots\right)$$

$$\epsilon^a-\epsilon^{-a}=2a\left(1+\frac{a^2}{\underline{|3}}+\frac{a^4}{\underline{|5}}+\ldots\right)$$

$$\frac{\epsilon^a+\epsilon^{-a}}{\epsilon^a-\epsilon^{-a}}=\frac{1+\frac{a^2}{\underline{|2}}+\frac{a^4}{\underline{|4}}+\ldots}{a\left(1+\frac{a^2}{\underline{|3}}+\frac{a^4}{\underline{|5}}+\ldots\right)}$$

$$\frac{I}{Nm}=\frac{1}{a}\left(1+\frac{a^2}{\underline{|2}}\right)\left(1+\frac{a^2}{\underline{|3}}\right)^{-1}-\frac{1}{a}$$

neglecting a^4, a^6, etc.

Then,
$$\frac{I}{Nm}=\frac{1}{a}\left(1+\frac{a^2}{\underline{|2}}\right)\left(1-\frac{a^2}{\underline{|3}}\right)-\frac{1}{a}$$

$$=\frac{1}{a}\left(\frac{a^2}{\underline{|2}}-\frac{a^2}{\underline{|3}}\right)$$

$$=\frac{a}{3}.$$

Since $a=\frac{mH}{kT}$, it follows that for low values of $\frac{H}{T}$, $\frac{I}{Nm}=\frac{mH}{3kT}$, or $I=\frac{m^2NH}{3kT}$. Thus the magnetic susceptibility $\kappa=\frac{I}{H}=\frac{m^2N}{3kT}$ is inversely proportional to the absolute temperature. This is in accord with experimental results (p. 564).

On the other hand, if a is not small,

$$
\begin{aligned}
\frac{I}{Nm} &= \frac{\epsilon^a+\epsilon^{-a}}{\epsilon^a-\epsilon^{-a}}-\frac{1}{a} \\
&= \frac{1+\epsilon^{-2a}}{1-\epsilon^{-2a}}-\frac{1}{a} \\
&= (1+\epsilon^{-2a})(1-\epsilon^{-2a})^{-1}-\frac{1}{a} \\
&= (1+\epsilon^{-2a})(1+\epsilon^{-2a}+\epsilon^{-4a}+\ .\ .\ .\ )-\frac{1}{a} \\
&= 1+2\epsilon^{-2a}+2\epsilon^{-4a}+\ .\ .\ .\ -\frac{1}{a}.
\end{aligned}
$$

As a increases, the right-hand side of this equation approaches unity, so that for very strong magnetic fields $\frac{I}{Nm}=1$, or $I=Nm$. This corresponds to the condition of saturation when all the atomic magnetic moments have the same direction as the magnetising field. Calling then the saturation value of the intensity of magnetisation I_0,

$$
\begin{aligned}
\frac{I}{I_0} &= \coth a-\frac{1}{a} \\
&= 1+2\epsilon^{-2a}+2\epsilon^{-4a}+\ .\ .\ .\ -\frac{1}{a}.
\end{aligned}
$$

The curve in Fig. 437 gives the manner in which $\frac{I}{I_0}$ increases with a, and if the student will plot it, it will be seen that $\frac{I}{I_0}$ approaches unity as a approaches infinity. Some idea of the magnitude of a may be obtained if the saturation value Nm were known. This condition cannot be even approached with any magnetic fields that can be produced artificially, but if we accept for a moment the value 1610 for iron (p. 289) then,

$$a=\frac{mH}{kT}=\frac{NmH}{NkT}=\frac{NmH}{RT}.$$

Now $R=8\cdot3\times10^7$ ergs per degree (p. 192) for 1 gramme-molecule of a gas, and taking the atomic weight of iron as 56 and its density 7·8, the volume of a gramme-atom is $\frac{56}{7\cdot8}$ c.c. compared with 1 c.c. for which R is calculated.

Therefore
$$a=\frac{56\times1610}{7\cdot8\times8\cdot3\times10^7}\frac{H}{T}$$
$$=1\cdot4\times10^{-4}\left(\frac{H}{T}\right).$$

At ordinary temperatures a is thus a small quantity and $\frac{I}{Nm}$ is small, and only the nearly straight part of the curve in Fig. 437 is realised for any practicable magnetic fields. The above theory applies properly to the case of a gas where the molecules can rotate without restriction, but it forms an approximate solution

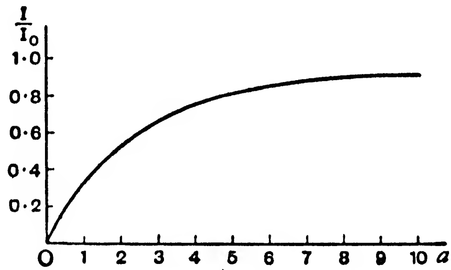


FIG. 437.

to the case of a paramagnetic liquid and to some extent to that of a solid.

Molecular Field.—It is to Weiss that the theory of molecular field is due. Ewing's theory, that the constraint on the elementary magnets (p. 294) is due to their neighbour's magnetic effect is now developed with the help of Langevin's method. In the interior of a magnetic body there are two fields, one, H, due to external causes and called the applied field, and another due to the magnetisation of the material. The resultant field is thus H+NI where N is some coefficient for each material. On p. 271 the field NI was due to poles on the surface of bodies, but in the present discussion it is due to neighbouring elementary magnets. Writing $H'=H+NI$

Then,
$$a=\frac{mH'}{kT} \quad . \; . \; . \quad \text{(p. 567)}$$

and,
$$\frac{I}{nm}=\coth a-\frac{1}{a},$$

where n is now taken for the number of molecules per c.c., to avoid confusion with N, above.

If M is the molecular weight of the substance and ρ its density, $\frac{M}{\rho}$ is the volume of one gramme-molecule, and $\frac{M}{\rho}I$ its resultant magnetic moment. This quantity is called the gramme-molecular magnetic moment σ; thus $\sigma=\frac{MI}{\rho}$.

Now,
$$I=\frac{H'-H}{N}$$

$$\therefore \quad \sigma=\frac{MH'}{\rho N}-\frac{MH}{\rho N}$$
$$=\frac{kTaM}{m\rho N}-\frac{MH}{\rho N}$$
$$=\frac{RTaM}{nm\rho N}-\frac{MH}{\rho N}$$

since $k=\frac{R}{n}$.

For saturation, all the elementary magnets are aligned with the magnetising field, and the number in one gramme-molecule is $\frac{nM}{\rho}$, so that the saturation gramme-molecular moment is $\frac{nMm}{\rho}=\sigma_0$, or $nm=\frac{\sigma_0\rho}{M}$

$$\therefore \sigma=\frac{RTaM^2}{\rho^2N\sigma_0}-\frac{MH}{\rho N}$$
$$\frac{\sigma}{\sigma_0}=\frac{RTaM^2}{\rho^2N\sigma_0{}^2}-\frac{MH}{\rho N\sigma_0}.$$

For any magnetic field which is not extremely great $\frac{I}{I_0}=\frac{\sigma}{\sigma_0}=\frac{a}{3}$, (p. 569).

$$\therefore \frac{\sigma}{\sigma_0}=\frac{3RTM^2}{\rho^2N\sigma^2}\cdot\frac{\sigma}{\sigma_0}-\frac{MH}{\rho N\sigma_0}$$
$$\frac{\sigma}{\sigma_0}\left(\frac{3RTM^2}{\rho^2N\sigma_0{}^2}-1\right)=\frac{MH}{\rho N\sigma_0}.$$

The magnetic susceptibility reckoned per gramme-molecule χ_M is $\frac{\sigma}{H}$

$$\therefore\ \chi_M=\frac{M}{\left(\dfrac{3RTM^2}{\rho^2 N\sigma_0{}^2}-1\right)\rho N}$$

$$=\frac{M}{\dfrac{3RM^2}{\rho^2 N\sigma_0{}^2}\left(T-\dfrac{\rho^2 N\sigma_0{}^2}{3RM^2}\right)\rho N}$$

$$=\frac{\dfrac{\sigma_0{}^2}{3R}\cdot\dfrac{\rho}{M}}{T-\dfrac{\rho^2 N\sigma_0{}^2}{3RM^2}}.$$

R has been taken as the gas constant for unit volume of substance. If, on the other hand, one gramme-molecule is taken, the corresponding constant R_M is $R\,.\,\dfrac{M}{\rho}$.

$$\therefore\ \chi_M=\frac{\dfrac{\sigma_0{}^2}{3R_M}}{T-\dfrac{\rho N\sigma_0{}^2}{3MR_M}}$$

Thus $\dfrac{1}{\chi_M}=K(T-\theta)$ where K and θ are constants. This is known as the Curie-Weiss law and is more general than Curie's law (p. 564). The temperature $\theta=\dfrac{\rho N\sigma_0{}^2}{3MR_M}$ is important because at temperatures below it a substance may be spontaneously magnetised.

Thus in the equation $$\frac{\sigma}{\sigma_0}=\frac{R_M TMa}{\rho N\sigma_0{}^2}-\frac{MH}{\rho N\sigma_0}$$

if H=0, $$\frac{\sigma}{\sigma_0}=\frac{R_M TMa}{\rho N\sigma_0{}^2}=\frac{T}{\theta}\cdot\frac{a}{3}.$$

This equation is plotted in Fig. 438 and gives the straight line OP whose slope is $\dfrac{T}{3\theta}$. Since this line cuts the curve at O and P, these points represent solutions of the equation giving OPQ and that giving the straight line. Thus $\dfrac{\sigma}{\sigma_0}$ may have the value zero or PS. If, however, a small increase in σ or a beyond the value PS or OS can be produced, then according to the straight line $\dfrac{\sigma}{\sigma_0}=\dfrac{T}{3\theta}a$, a point above the equilibrium curve would be reached, which means that N would be negative, and the molecular field would bring the magnetisation back to the equilibrium condition

represented by the curve. On the other hand, a small increase in σ or a at O would, according to the line $\frac{\sigma}{\sigma_0}=\frac{T}{3\theta}a$ produce a molecular field in the direction of approach towards the equilibrium curve, that is, in a direction to produce further magnetisation. This increase would in turn produce further magnetisation. In fact, the condition of magnetisation from O to P is unstable and the material would, in the absence of applied field, become spontaneously magnetised to an amount represented by P. There will be no spontaneous magnetisation unless the straight line OP lies below the curve and this only occurs if its slope is less than that of the curve at the origin. Now the slope of the curve at the origin is $\frac{a}{3}$ (p. 570) and the condition for spontaneous magnetisation is therefore $\frac{\sigma}{\sigma_0}<\frac{a}{3}$, or $\frac{MR_MT}{n\rho\sigma_0^2}a<\frac{a}{3}$, that is, $T<\frac{n\rho\sigma_0^2}{3MR_M}$. The latter quantity is θ, and it follows that $T<\theta$. Above the temperature θ the substance is paramagnetic, and this temperature is therefore a transition temperature.

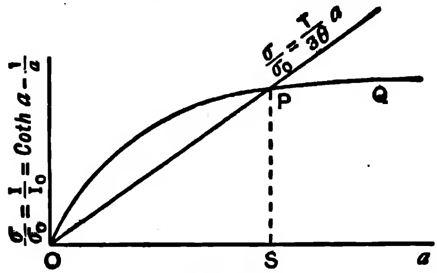


FIG. 438.

Ferromagnetism: Weiss.—According to Weiss, a ferromagnetic material is an agglomeration of regions, each of which is spontaneously magnetised, but the directions of spontaneous magnetisation are distributed indiscriminately in the material. Such regions or "domains" are known to exist (p. 295). An applied field in the direction of magnetisation will have no effect, but a reversed field, if sufficiently great, may cause a complete reversal of the magnetisation. One domain would thus have a simple rectangular hysteresis loop, but the overall effect of a great number disposed at random is to give the normally smooth experimental curves. However, to account for the reversible magnetisation which occurs for instance in weak fields (p. 287), it is necessary to suppose that the individual magnetic dipoles in the boundary layer between two domains can swing round into alignment with whichever domain has its magnetic direction closest to that of the applied field. In this way one domain can grow reversibly at the expense of a neighbour.

Weiss believed that he had discovered the magnitude of the moment of the fundamental magnet, the Weiss Magneton, from

the observation that at low temperatures the gramme-molecular moment σ_0 had the values 12,360 and 3370 respectively for iron and nickel, which are very nearly 11 and 3 times 1123·5. Dividing this by $6{\cdot}064\times10^{23}$, the number of molecules in a gramme-molecule, the magneton is found to have the value $1{\cdot}85\times10^{-21}$. A somewhat higher value, 1125·6, was later adopted for the gramme-molecular constant. Although there is now no reason to suppose that the Weiss magneton is fundamental, it remains a convenient unit.

Modern Theory.—Later in this chapter, where atomic theory and the theory of the solid state are discussed, it is explained that the essential feature of a paramagnetic atom or ion is now held to be the possession of a spinning electron (p. 615) unpaired by an electron of oppositely directed spin. The forces between electrons in these circumstances are such that a mutual alignment occurs only if the interatomic spacing is within quite narrow limits, and if this condition is obeyed, at sufficiently low temperatures ferromagnetism appears. Since this spacing depends essentially on the crystal structure, we see how it is possible for typically ferromagnetic ions such as iron to be present in a paramagnetic salt or alloy and yet for " non-magnetic " substances to form ferromagnetic alloys (p. 286).

Production of very low Temperature.—Owing to the connection between heat and magnetisation it has been suggested by Debye[1] and Giauque[2] that by the adiabatic demagnetisation of paramagnetic substances extremely low temperatures may be produced. The method is to cool the substance in liquid helium to a temperature 1·3° above absolute zero of temperature, that is, 1·3° K. It is then magnetised in as strong a field as possible and the resulting heat produced is allowed to leak away. On removing the field the substance becomes demagnetised adiabatically and very low temperature results. It is necessary to use a substance which has not abnormally high specific heat at these temperatures. Haas and Wiersma[3] using a mixture of potassium chrome alum and potassium aluminium alum obtained a temperature of 0·0034° K.

Temperature Equilibrium of Electrons.—We shall now consider briefly the theory of electrical conduction by means of free electrons, first suggested by Riecke[4] and more explicitly stated by Drude.[5] Remembering the distinction we have drawn

[1] P. Debye, *Ann. der Phys.*, **81**, p. 1154. 1926.
[2] W. F. Giauque, *J. Am. Chem. Soc.*, **49**, p. 1864. 1927.
[3] W. J. de Haas and E. C. Wiersma, *Physica*, **2**, p. 335 and p. 438. 1935.
[4] E. Riecke, *Wied. Ann.*, **66**, pp. 353 and 545. 1898.
[5] P. Drude, *Ann. der Physik.*, **1**, p. 566 (1900); and **3**, p. 369 (1900).

between electrical conductors and insulators, that in the latter the electrons are bound within the atom, whereas in the former some of them are free to leave the atom and therefore move under forces exerted by an electric field, we may, by assuming that these free ions obey the ordinary laws found for molecules in the kinetic theory of gases, obtain certain results which are in agreement with experimentally observed facts.

Let us consider for a moment the elementary form of the kinetic theory of gases. If a gas consist of N molecules per unit volume, which are entirely independent of each other, except that in their motions they will frequently collide, we may, as a first approximation, resolve all the molecular motions parallel to the three rectangular axes, and remembering that at any instant as many are moving parallel to an axis in its positive direction as in its negative direction, we can represent the indiscriminate motions of the molecules by dividing them into six groups of constant velocity, each moving in one particular direction parallel to one of the three axis. Thus we have $\frac{N}{6}$ molecules moving in one direction with say velocity v. A plane surface of unit area at right angles to this direction will therefore be struck by $\frac{Nv}{6}$ molecules per second. Each molecule has momentum $+mv$ before striking the surface and momentum $-mv$ on rebounding from it, so that the impulse at each rebound is $2mv$. Hence the pressure on the surface is $\frac{Nv}{6} \times 2mv$,

or,
$$p = \tfrac{1}{3}Nmv^2$$

Nm is the mass per unit volume of the gas, or its density, and since pressure × volume = RT, where T is the absolute temperature, we see, since the volume is unity, that—

$$\tfrac{1}{3}Nmv^2 = RT.$$

If 1 gramme-molecule of the gas be taken, R is the universal gas constant, or $8{\cdot}32 \times 10^7$ (p. 192), m is the mass of one molecule, and N the number of molecules in unit volume, which is the same for all gases at the same temperature and pressure.

$$\therefore \tfrac{1}{3}mv^2 = \frac{R}{N}T,$$

or,
$$\tfrac{1}{2}mv^2 = \frac{3}{2} \cdot \frac{R}{N}T.$$

Hence we conclude that for all gases in equilibrium, the abso-

lute temperature is proportional to the mean kinetic energy of the molecules,

or, $$\tfrac{1}{2}mv^2 = \alpha T, \text{ where, } \alpha = \frac{3}{2} \cdot \frac{R}{N}.$$

Now, the most probable value of the ionic charge e is $1{\cdot}59 \times 10^{-20}$ electromagnetic unit, and therefore the passage of 1 electromagnetic unit of charge through acidulated water liberates $\frac{1}{1{\cdot}59 \times 10^{-20}}$ ions at each electrode. But this also liberates 0·0001045 gramme of hydrogen, and therefore the number of atoms per gramme of hydrogen is $\frac{10^{24}}{1{\cdot}59 \times 1{\cdot}045}$, and since one gramme-molecule of hydrogen is 2 grammes, the number of atoms to the molecule being 2, the number of molecules to the gramme-molecule, or N

$$= \frac{10^{24}}{1{\cdot}59 \times 1{\cdot}045} = 6{\cdot}03 \times 10^{23},$$

$$\therefore \ \alpha = \frac{3}{2} \times \frac{8{\cdot}315 \times 10^7}{6{\cdot}03 \times 10^{23}} = 2{\cdot}07 \times 10^{-16}.$$

In the case of hydrogen at 0° C. and 76 cm. pressure we may easily find the mean square velocity of the molecules, for

$$p = \tfrac{1}{3} N m v^2 = \tfrac{1}{3}(\text{density}) v^2,$$

$$\therefore \ 76 \times 13{\cdot}596 \times 981 = \tfrac{1}{3}(0{\cdot}00008987) v^2,$$

from which, $$v^2 = 3{\cdot}38 \times 10^{10},$$

and, $$v = 1{\cdot}14 \times 10^5 \text{ cm. per sec.}$$

If now the electrons within a conductor are in temperature equilibrium with their surroundings, we can find their velocity according to the kinetic theory; for $\frac{1}{2}mv^2$ is the same for them as for any other gas molecules at the same temperature. Taking their mass as $\frac{1}{1835}$ of that of a hydrogen atom,

$$\tfrac{1}{2} m_h v_h^2 (\text{for hydrogen}) = \tfrac{1}{2} m_e v_e^2 (\text{for electron}),$$

$$\therefore \ v_e^2 = \frac{m_h}{m_e} \cdot 3{\cdot}38 \times 10^{10}$$

$$= 1835 \times 3{\cdot}38 \times 10^{10}$$

$$= 6{\cdot}20 \times 10^{13}$$

$$v_e = 7{\cdot}88 \times 10^6 \text{ cm. per sec.}$$

Electrical Conduction.—In addition to this motion of the electrons which takes place in all directions, we shall have, in the presence of an electric field, a drift in the direction of the field, or rather in the opposite direction, since the charges are negative.

Assuming that after each collision with a molecule of the conductor, the electron starts afresh with the velocity corresponding to the temperature equilibrium with the substance, it will in the time that elapses before the next collision, be subject to the influence of the electric field, and will on that account have an acceleration $f=\frac{Ee}{m}$ parallel to the field, where E is the electric intensity.

The different electrons will have various distances of travel between two collisions, but the average distance is called the length of mean free path, λ, and the time taken in describing this mean free path is $\frac{\lambda}{v}$. Hence the distance travelled in the direction of the field on account of the acceleration $\frac{Ee}{m}$, is $\frac{1}{2}\cdot\frac{Ee}{m}\cdot\frac{\lambda^2}{v^2}$, and the average velocity in this direction is—

$$\frac{\text{distance}}{\text{time}}=\frac{1}{2}\cdot\frac{Ee}{m}\cdot\frac{\lambda^2}{v^2}\cdot\frac{v}{\lambda}$$

$$=\frac{1}{2}\cdot\frac{Ee\lambda}{mv}.$$

But, $$\tfrac{1}{2}mv^2=\alpha T,\text{ or, } m=\frac{2\alpha T}{v^2},$$

$$\therefore\ \text{average velocity in direction of field}=\frac{Ee\lambda v}{4\alpha T}.$$

The corresponding current is $Ne\times$velocity, where N is the number of electrons per unit volume,

$$\therefore\ \text{current}=\frac{NEe^2\lambda v}{4\alpha T}=i,$$

and conductivity, $$\sigma=\frac{i}{E}=\frac{Ne^2\lambda v}{4\alpha T}=\frac{Ne^2\lambda}{4\alpha T}\sqrt{\frac{2\alpha T}{m}}=\frac{Ne^2\lambda}{2\sqrt{2\alpha Tm}}.$$

Hence Ohm's law is obeyed, for the conductivity is independent of the current, but the conductivity should be inversely proportional to the square root of the absolute temperature. For most pure metals the conductivity is found to vary inversely as the absolute temperature, or, what is the same thing, the resistivity is proportional to the absolute temperature. The curves for some of the pure metals, taken from Dewar and Fleming's [1] results, show that down to quite low temperatures the resistivity behaves as though it would vanish at the absolute zero of temperature (Fig. 439).

In the above expression for the electrical conductivity, the

[1] J. Dewar and J. A. Fleming, *Phil. Mag.*, 36, p. 271. 1893.

only quantities that are likely to vary from one substance to another are N and λ. Now, it is improbable that λ would vary greatly, so that it is concluded that the great differences in conductivity between different materials arise from the difference in the quantities of free electrons in them.

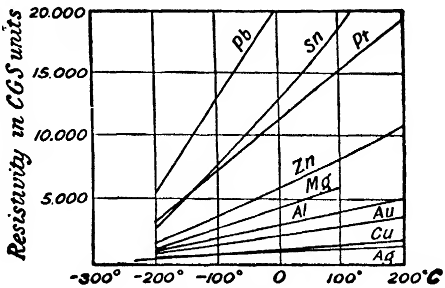


FIG. 439.

Superconductivity.—The development of low-temperature research, made possible by the production of liquid helium, has brought to light the fact that many metals lose almost all their resistivity at low temperatures. Kamerlingh Onnes [1] found that solid mercury immersed in liquid helium lost its resistivity, or became superconducting at a temperature of about 4·2° K., that is, 4·2° above absolute zero.

Several methods have been employed for exhibiting and measuring superconductivity. The fall of potential across the conductor may be compared with that across a standard resistance when the same current flows in both. Again, a suspended magnet may be placed externally to a superconducting coil in which a current has been started, when the deflection of the magnet may last for many hours or days since the time constant of the coil (L/R) (p. 309) is so great. Or again, a form of dynamometer made of two coils of the superconducting material exhibits a couple acting on the suspended coil for a considerable time when a current has been started by means of a magnetic field. With this last form, using rings of lead, the current did not decrease in one hour by as much as 1 in 40,000.

The transition does not, in most cases, occur suddenly but may be distributed over several hundredths of a degree. The crystalline state influences the rate of transition, single crystals having the sharpest change. The transition temperature is taken to be midway between the end and beginning of transition. For several common substances it is—Pb, 7·26° K.; Hg, 4·12° K.; Sn, 3·69° K.; Al, 1·14° K.; Zn, 0·79° K., and Cd, 0·6° K.

[1] Kamerlingh Onnes, *Leiden Comm.* Nos. 122*b*, 124*c*, 133*a*, 133*c*.

A peculiarity of superconductivity is that in fairly high magnetic fields it disappears, the normal conductivity returning. This occurs for Pb at 600 gauss at 4·2° K., Sn 50 gauss at 3·35° K., Hg 14 gauss at 4·12° K.

There is, as yet, no satisfactory explanation of superconductivity.

Thermal Conductivity.—The close connection between electrical and thermal conductivity has long been known; in fact, the law of Wiedemann and Franz [1] states that the ratio of the conductivity for heat to the electrical conductivity at any temperature is constant for all metals, and is proportional to the absolute temperature. Drude's theory gives an explanation of this, by treating the electrons in the conductor in a manner similar to that in which molecules are dealt with in the kinetic theory of gases. Consider a plane AB in the conductor, and two others E and F parallel to it and at distances from it equal to the mean free path of the electrons (Fig. 440). Returning to our method of dividing the total number of electrons N per unit volume into six groups, we see that $\frac{N}{6}$ are crossing unit area of AB from left to right, while an equal number cross from right to left. The remaining groups are moving parallel to AB and will not further concern us. If the temperature is the same everywhere, those passing from left to right carry the same amount of energy as those from right to left, and the resultant energy transferred in any direction is zero.

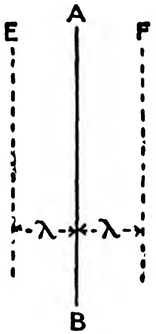


FIG. 440.

If now the temperature at the plane E is T_1 and is higher than the temperature T_2 at F, those passing from left to right have greater energy than those passing in the opposite direction, and there is a transference of heat through AB from left to right. Only those electrons that lie between E and AB will cross without collision, since any on the left of E are at a distance greater than the mean free path and will experience collision without reaching AB, and their direction then becomes changed.

Again, the number of electrons per unit volume travelling from left to right is $\frac{N}{6}$, and since their velocity when passing AB is v, the number crossing unit area of AB in unit time is $\frac{N}{6}v$, and since each has kinetic energy $\frac{1}{2}mv_1^2$, where v_1 is the velocity corresponding to temperature T_1 at E,

[1] G. Wiedemann and R. Franz, *Pogg. Ann.*, 89, p. 497. 1853.

Energy carried in one second from left to right through unit area is

$$\frac{N}{6}v \cdot \tfrac{1}{2}mv_1^2 = \frac{N}{6}va\mathrm{T}_1 \quad (\text{since } \tfrac{1}{2}mv_1^2 = a\mathrm{T}_1).$$

Similarly, energy passing from right to left $= \frac{N}{6}va\mathrm{T}_2$.

Therefore, balance of energy carried in one second through unit area of AB $= \frac{N}{6}va(\mathrm{T}_1 - \mathrm{T}_2)$.

It should be noticed that the number crossing in each direction must be the same, since the pressure at all points remains constant. If this were not the case there would be a drift of electrons in one direction or the other, and the process of transference of heat would not be one of pure conduction.

Now, if the thermal conductivity be k, then, since the temperature gradient at AB is $\frac{\mathrm{T}_1 - \mathrm{T}_2}{2\lambda}$, the transference of heat per second through unit area is $k \times \frac{\mathrm{T}_1 - \mathrm{T}_2}{2\lambda}$.

Hence,
$$\frac{N}{6}v \cdot a(\mathrm{T}_1 - \mathrm{T}_2) = k\frac{\mathrm{T}_1 - \mathrm{T}_2}{2\lambda},$$
$$\therefore\ k = \frac{Nva\lambda}{3}.$$

Remembering that the electrical conductivity is given on p. 578 by

$$\sigma = \frac{Ne^2\lambda v}{4a\mathrm{T}},$$

we see that,
$$\frac{k}{\sigma} = \left(\frac{a}{e}\right)^2 \cdot \tfrac{4}{3}\mathrm{T},$$

which is in agreement with the law of Wiedemann and Franz.

	$\frac{k}{\sigma}$	Temperature coefficient.
Copper	$6{\cdot}71 \times 10^{10}$	$3{\cdot}95 \times 10^{-3}$
Silver	$6{\cdot}86 \times 10^{10}$	$3{\cdot}77 \times 10^{-3}$
Gold	$7{\cdot}09 \times 10^{10}$	$3{\cdot}75 \times 10^{-3}$
Zinc	$6{\cdot}72 \times 10^{10}$	$3{\cdot}85 \times 10^{-3}$
Lead	$7{\cdot}15 \times 10^{10}$	$4{\cdot}07 \times 10^{-3}$
Tin	$7{\cdot}35 \times 10^{10}$	$3{\cdot}4 \times 10^{-3}$
Platinum	$7{\cdot}53 \times 10^{10}$	$4{\cdot}64 \times 10^{-3}$
Bismuth	$9{\cdot}64 \times 10^{10}$	$1{\cdot}51 \times 10^{-3}$
Constantan	$11{\cdot}06 \times 10^{10}$	$2{\cdot}39 \times 10^{-3}$
Manganin	$9{\cdot}14 \times 10^{10}$	$2{\cdot}74 \times 10^{-3}$

The table on p. 581 gives a few of the ratios $\frac{k}{\sigma}$ determined by Jäger and Diesselhorst.[1]

It will be seen that for the pure metals, with the exception of bismuth, the ratio $\frac{k}{\sigma}$ is very fairly constant, and the temperature coefficient is nearly $3{\cdot}66 \times 10^{-3}$, the coefficient of expansion of a gas. Hence the electronic theory gives some approximation to a true explanation of these processes.

Again, since $\alpha = 2{\cdot}04 \times 10^{-16}$, and $e = 1{\cdot}59 \times 10^{-20}$, we have—

$$\frac{k}{\sigma} = \left(\frac{2{\cdot}04 \times 10^{-16}}{1{\cdot}59 \times 10^{-20}}\right)^2 \times \tfrac{1}{3} \times 273$$
$$= 5{\cdot}99 \times 10^{10}.$$

This value is sufficiently near the observed value to strengthen the theory considerably, especially in view of the simplifying assumptions which we have made.

The electron theory of conduction has been considerably modified by the application of the quantum theory. A short and much simplified version of the modern treatment of electrical conduction is given later in this chapter (p. 616).

Emission of Electricity by Hot Bodies.—It has been known for a long time that an electric charge would leak much more rapidly from a hot body than from a cold one, and that the rate of leak is different for charges of opposite signs. The phenomenon was investigated by Elster and Geitel.[2] A wire is heated by means of a current, and a plate near it is connected to an electrometer. The charge received by the plate depends upon the gas present, its pressure, and also upon the nature of the wire. The temperature of the wire, however, is the most important factor in determining the electrification of the plate. With oxygen at atmospheric pressure, and the wire at a dull red heat, the plate receives a positive charge, its potential being 2 or 3 volts. With rise of temperature the charge increases, reaches a maximum, and then falls to a very low value. Reduction in the pressure to that of a high vacuum reduces the charge, and even reverses its sign at very low pressures. With hydrogen the charge on the plate is negative at all pressures. The effects are exceedingly complicated, as the wire gives out occluded gas and may even give off pieces of its own substance.

Harker and Kaye[3] have obtained currents as great as 10 amperes by means of an electromotive force of 8 volts between

[1] W. Jäger and H. Diesselhorst, *Wiss. Abh. der Phys. Tech. Reichsanstalt*, **8**, 1900.
[2] J. Elster and H. Geitel, *Wied. Ann.*, 1882, 1883, 1884, 1885, 1887 and 1889.
[3] J. A. Harker and G. W. C. Kaye. *Proc. Roy. Soc.*, A. **86**, p. 379. 1912.

carbon rods at a temperature approaching 3000° C., the pressure being atmospheric. Between carbon rods at different temperatures they have obtained currents on account of the different rates of emission of ions.

Thermionics.—The phenomenon of the emission of electricity from hot bodies has been subjected to extended investigation by Prof. Richardson,[1] who gave the name *thermionics* to the subject and *thermions* to the ions emitted. The general method of experiment is to heat the substance, in the form of a wire, by means of an electric current. The wire is situated inside a metal cylinder in a very highly evacuated tube, and the current between the wire and the cylinder is measured by means of a capacity and electrometer (p. 514) for small currents and by means of a galvanometer for larger currents. Measurement of the resistance of the wire determines its temperature. It is found that at high vacuum, when gaseous contamination is eliminated, the emission of electrons is very regular and increases rapidly with the temperature. Applying the gas laws to the electrons in equilibrium with the metal at any temperature, Richardson found from thermodynamic reasoning that the ionisation current from the metal is of the form

$$i = \mathrm{AT}^{\frac{1}{2}} e^{-\frac{b}{\mathrm{T}}}.$$

For, let an enclosure with a piston, and one side consisting of the metal, be taken round a reversible cycle as described for the reversible cell on p. 187. The cycle is started with an infinitesimal adiabatic expansion from T to T$-\delta$T, followed by an isothermal compression in which nv electrons are driven into the metal. Then there is an adiabatic compression followed by an isothermal expansion in which nv electrons pass from the metal to the space, n being the number of electrons per unit volume of space in equilibrium with the metal at temperature T. No heat is absorbed or given out during the adiabatic changes. The resultant work done is then $v\delta p$, since the number of electrons in the space at the finish of the cycle is the same as at the beginning. Also the work done during the isothermal change at T is $n\phi v + pv$, where ϕ is the work required to remove one electron from the metal to the space; then, from the second law of thermodynamics,

$$\frac{\text{work for cycle}}{\text{work absorbed at T}^\circ} = \frac{\delta \mathrm{T}}{\mathrm{T}};$$

$$\therefore \frac{v\delta p}{n\phi v + pv} = \frac{\delta \mathrm{T}}{\mathrm{T}},$$

or,
$$\mathrm{T}\delta p = n\phi\delta\mathrm{T} + p\delta\mathrm{T}.$$

[1] O. W. Richardson, "The Emission of Electrons from Hot Bodies." Longmans, Green & Co. 1922.

If the electrons in the space be considered to constitute a gaseous state, the gas equation applies to them ; that is,

$$p=nk\mathrm{T},$$

where k is the gas constant expressed in terms of one electron,

$$\delta p=nk\delta\mathrm{T}+\mathrm{T}k\delta n$$

$$\therefore\ nk\mathrm{T}\delta\mathrm{T}+k\mathrm{T}^2\delta n=n\phi\delta\mathrm{T}+nk\mathrm{T}\delta\mathrm{T}$$

$$\frac{\delta n}{n}=\frac{\phi}{k\mathrm{T}^2}\delta\mathrm{T}.$$

Integrating, we have, $$\log n=\int\frac{\phi}{k\mathrm{T}^2}d\mathrm{T}+\mathrm{C}$$

or, $$n=\mathrm{A}\epsilon^{\int\frac{\phi}{k\mathrm{T}^2}d\mathrm{T}}.$$

From the kinetic theory of gases, the number of electrons meeting unit area of the boundary per second is $n\sqrt{\frac{k\mathrm{T}}{2\pi m}}$, where m is the mass of an electron, and when equilibrium is reached this is equal to the rate of emission, which is therefore $\mathrm{A}\sqrt{\frac{k}{2\pi m}}\mathrm{T}^{\frac{1}{2}}\epsilon^{\int\frac{\phi}{k\mathrm{T}^2}d\mathrm{T}}$, and the thermionic current from the metal is therefore e times this. If ϕ is considered to be independent of T,

$$\int\frac{\phi}{k\mathrm{T}^2}d\mathrm{T}=-\frac{\phi}{k\mathrm{T}},$$

and, $$i=\mathrm{A}\mathrm{T}^{\frac{1}{2}}\epsilon^{-\frac{b}{\mathrm{T}}}$$

where A and b are constants depending upon the metal. Richardson gives a thermodynamic reason for considering that $\phi=\phi_0+\frac{3}{2}k\mathrm{T}$, in which case,

$$\int\frac{\phi}{k\mathrm{T}^2}d\mathrm{T}=\int\frac{\phi_0}{k\mathrm{T}^2}d\mathrm{T}+\frac{3}{2}\int\frac{d\mathrm{T}}{\mathrm{T}}$$

$$=-\frac{\phi_0}{k\mathrm{T}}+\frac{3}{2}\log\mathrm{T};$$

and $$\epsilon^{\int\frac{\phi}{k\mathrm{T}^2}d\mathrm{T}}=\epsilon^{-\frac{\phi}{k\mathrm{T}}}\epsilon^{\frac{3}{2}\log\mathrm{T}}$$

$$=\epsilon^{-\frac{\phi}{k\mathrm{T}}}\mathrm{T}^{\frac{3}{2}},$$

so that, $$i=\mathrm{A}\mathrm{T}^2\epsilon^{-\frac{b}{\mathrm{T}}}.$$

Both these equations represent the thermonic current very well, but the latter is probably the more correct, as the relation between ϕ and T has not been assumed. Taking the relation $i=\mathrm{A}_1\mathrm{T}^{\frac{1}{2}}\epsilon^{-\frac{b}{\mathrm{T}}}$, Richardson has found that for platinum

$$\mathrm{A}_1=7{\cdot}5\times10^{25},\quad b=4{\cdot}93\times10^4,\quad\text{and}\quad\phi_0=4{\cdot}1\text{ volts.}$$

Richardson also applies the equilibrium condition of electrons near the metals to explain the contact difference of potentials between metals and the Peltier and Thomson effects. For an account of this and the effect of the presence of gases upon the metals and the application of the quantum theory to the phenomenon, the student is referred to Richardson's book (p. 583).

Edison Effect.—The Edison effect, made use of by Sir A. Fleming in his oscillation valve (p. 452), is another result of the same phenomenon. If a metallic plate, D, be situated between the limbs of the filament of an incandescent lamp, then on connecting the positive end of the filament through a galvanometer to D (Fig. 441) a current will be observed to flow in the galvanometer, but none when D is connected to the negative limb B. We should expect that the negative ions emitted by the incandescent carbon would be more vigorously repelled from the negative limb, and hence on meeting D would lower its potential. The difference of potential between A and D is then much greater than that between B and D.

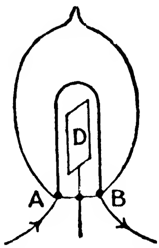


FIG. 441.

It has been shown by Sir J. J. Thomson,[1] by the method of finding the effect of a magnetic field upon the rate of leakage (p. 473), that the negative ions emitted by a hot wire have the same value of $\frac{m}{e}$ as the corpuscles in the cathode rays, thus establishing their identity with these bodies.

The emission of positive ions at temperatures below that required for the emission of negative ions has already been noticed, and their ratio $\frac{m}{e}$ was also found by observing the strength of magnetic field required to affect the leakage of a charge to a neighbouring conductor, and it was found that there were two sets of positive ions taking part in the leak, one set having a mass equal to that of the atom of the metal and the other that of the gas. In some cases there are bodies of greater mass still, taking part in the production of leakage. These are probably metallic dust.

Amplifying and Rectifying Valves.—The principle of the Fleming valve (p. 452) has been developed in recent years into a method for magnifying or amplifying extremely feeble electrical oscillations, by employing a vacuum bulb having three electrodes known as a triode. This amplification has greatly extended the range of radio-telegraphy and has also many other applications.

[1] J. J. Thomson, *Phil. Mag.*, 48, p. 547. 1899.

There are several forms of the valve, but a typical specimen has a filament of tungsten, F (Fig. 442), heated by a current from a battery of two cells BD, a gauze or spiral wire G, and a metal sheath P which is outside the gauze. If the bulb is very highly exhausted no current can flow from the filament to the gauze or grid G, through the valve, but current can flow in the reverse direction. The reason is that the hot filament emits electrons, which travel from the filament to the grid when the electric field is directed from the grid to the filament, but when the field is reversed the electrons remain on the filament. The circuit PTEDF contains the battery DE of 50 to 150 volts, which tends to drive a current through the valve from P to F, making P anode. The resistance of the valve between P and F is controlled greatly by the electric field between G and F, that is, by the potential of the grid with respect to the filament. For when G

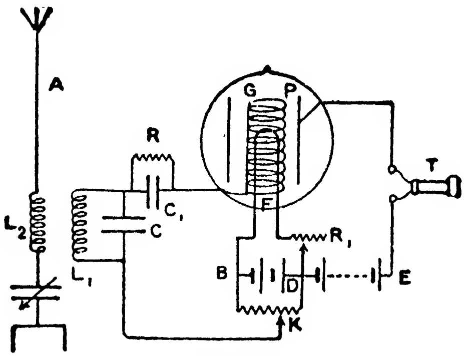


FIG. 442.

is negative with respect to F no electrons leave F, and the resistance of the valve is practically infinite. But if G is positive with respect to F, electrons pass freely from F and may pass in quantity through the interspaces of the grid, when there is an electric field from P to F. The current produced in the valve between the anode P and the filament depends, therefore, upon the p.d. between the grid and filament, and in order to understand the action of the valve, it is desirable to measure the grid current and anode current for various values of the grid-filament p.d. This may be done by replacing the telephone T by a micro-ammeter and placing a voltmeter between B and G, and a micro-ammeter in the circuit BCG. On varying the grid potential by varying the position of the potentiometer contact K, the three quantities are observed, and their values may be plotted. The curve obtained is called the static characteristic of the valve. The characteristic varies with different kinds of valve, but it is always of the same type in the case of hard

valves (Fig. 443), or valves where the vacuum is so high that there is no appreciable effect due to contained gas. The anode current is much greater than the grid current, and the latter is drawn to a larger scale than the former. Consequently the valve acts as an amplifier of small oscillatory current occurring in the circuit L_1C (Fig. 442).

In order to use the valve as a rectifier in radio-telegraphy, many devices are used. In the method considered here K is connected to D and a very small condenser C_1 is placed in the grid circuit, the condenser having a high resistance R, of the order of a megohm, in parallel with it. When there are no oscillations in the circuit L_1C, the grid potential acquires a steady value n (Fig. 443), a small current, nf, flowing from grid to filament. When electro-magnetic waves fall upon the aerial, oscillations of potential, na and nb, are produced in the grid circuit, and owing to the curvature of the grid current curve the increase of current ec is greater than the decrease fg, as in the

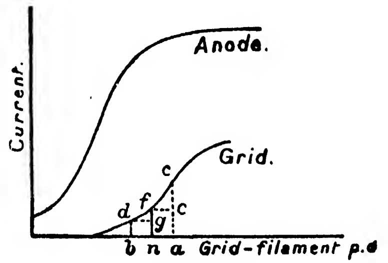


FIG. 443.

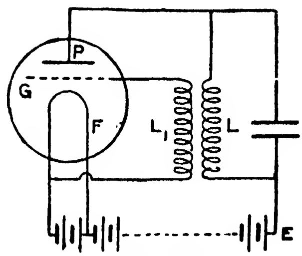


FIG. 444.

case of the crystal detector (p. 452). This means that the stream of electrons from the filament to the grid is increased, while a train of waves is arriving, and the potential of the grid is lowered. The anode current therefore suffers a decrease, but recovers its value corresponding to the point n when the waves cease. This change of mean current in the anode circuit causes a sound in the telephone. In this way the waves which are of too high a frequency to affect the telephone, or to be audible if they did, are enabled to produce signals which can be heard.

These three-electrode valves have also been used for generating continuous electromagnetic waves. On coupling the grid and anode circuits by means of a mutual inductance or by means of a capacity, and applying the battery to the anode circuit, it is found that, with suitable grid current, oscillations occur in the anode circuit. One method of carrying this out is shown in Fig. 444, in which the inductance L_1 is placed in the grid circuit and the oscillatory circuit LC is placed between P and E. The theory of production of the oscillations is not simple, but it may

be noted that for them to occur, the mutual inductance L_1L must be negative, so that when the current in L first rises the E.M.F. in L_1 raises the potential of G above that of F. The action of the valve is therefore to increase the anode current on rising and to diminish it when falling, so compensating for the damping which occurs in LC (p. 334), due to resistance.

In addition to the control grid present in the triode or three-electrode valve discussed above, a valve may have other grids. Thus a screen grid between the control grid and the anode, maintained at some intermediate potential, reduces the effect of the anode potential on the anode current and enhances the effect of changes in grid potential. Thus tetrodes (four electrodes) and pentodes (with yet another grid) in general amplify more than triodes.

Ionisation in Flames.—That the phenomenon of ionisation takes place in ordinary flames may be shown in several ways. For example, if two platinum wires are placed in a bunsen-burner flame but not touching each other, a current may be made to pass between them by connecting them to the terminals of a cell, and a sensitive galvanometer in the circuit will indicate a feeble current. If a bead of a sodium or potassium salt be placed in the flame below the platinum wires, the conductivity of the flame is enormously increased, owing to the presence of ions liberated from the substance at high temperature.

The old experiment of discharging an electrified glass or ebonite surface by passing a flame over it illustrates the presence of the ions, since those of opposite sign to the charge on the plate are attracted to it and neutralise the charge.

The increase in conductivity of a flame due to the introduction into it of a volatilisible metallic salt has been measured by several experimenters. Arrhenius [1] supplied the salt to the flame by spraying a solution into the gas which feeds the flame, and the concentration of the salt in the flame was determined by observing the rate at which a bead disappears which gives the same illumination as the spray. The conductivity in the flame is found by observing the current in a circuit which includes part of the flame, and subtracting the current produced when there is no salt employed. The (electromotive force)-current curve exhibits the same characteristics as that for an ionised gas, but the straight portion is not quite horizontal, showing that complete saturation is not attained. Using the same method, Prof. H. A. Wilson [2] found that for the salts of cæsium, rubidium, potassium, sodium, lithium and hydrogen, the conductivity is in the order of the atomic weights.

Prof. Wilson also found the velocity of the ions in a given

[1] S. Arrhenius, *Wied. Ann.*, **42**, p. 18. 1891.
[2] H. A. Wilson, *Phil. Trans.*, A, **192**, p. 499. 1899.

electrical field by arranging two electrodes in the flame, one above the other, with the bead of salt between them, and determining the field necessary to drive the ions downwards in opposition to their velocity due to the upward motion of the flame gases (Fig. 445). With the upper electrode positive, the presence of the bead will not affect the current unless the field is sufficiently strong to drive the positive ions downward with a velocity just greater than the velocity with which they are carried upwards by the flame. In this way the velocity of the negative ions for a potential gradient of one volt per cm. in a flame whose temperature is about 2000° C. was found to be about 1000 cm. per second. The velocities of the positive ions of the salts of cæsium, rubidium, potassium, sodium and lithium were all about 62 cm. per sec.

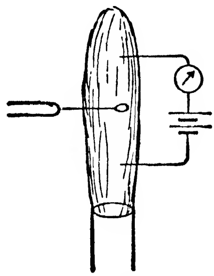


FIG. 445.

Using a stream of hot air at about 1000° C., the velocities were respectively 26 cm. per sec. for the negative, and 7·2 cm. per sec. for the positive ions, whereas for barium, strontium and calcium it is 3·8 cm. per sec. for the positive ions. These low velocities appear to indicate that the ions become loaded with neutral atoms, and the equality in velocities for the ions of the different atoms indicates that the size of these groups depends upon the charge on the ion, being larger in the case of the divalent ions than in the case of those which are monovalent.

Atoms.—It has already been seen that some of the most important properties of the atom depend upon its atomic number rather than its atomic weight. Thus Moseley's work on the characteristic X-rays (p. 496) proved an intimate connection between the frequency of the X-rays emitted and the atomic number, and Soddy and Fajans (p. 539) showed that the chemical properties of atoms depended upon their atomic number, not their atomic weight. Many attempts have been made to represent the constitution of the atom, some of these attempts being highly successful in giving an explanation of particular properties of matter. The explanations differ in point of view, according to the particular properties that are considered; but they concur in representing the atom as consisting of a nucleus surrounded by electrons. Some of these electrons are detachable from the atom; but in the ordinary neutral condition the total electrical charge of the atom is zero. It follows that the resultant positive charge associated with the nucleus is equal to the sum of the charges of the electrons surrounding it. Lord Rutherford [1] has shown from experiments on the scattering

[1] E. Rutherford, *Phil. Mag.*, 21, p. 669 (1911); and 37, p. 537 (1919).

of α rays on passing through matter, that the positive nucleus of the atom causing the scattering of the α rays is an extremely small body, and that the laws of force between the α particle and the nucleus of the atom which would cause the observed deflection of the former is the inverse square law. Further, the suggestion was made by Van der Broek[1] that the positive charge in the nucleus, measured in electronic units, is equal to the atomic number of the kind of atom considered, which result is now confirmed by Rutherford's work on the scattering of α rays. These considerations suggest that the hydrogen atom is a nucleus with one electron revolving around it; a helium atom is a nucleus of two positive units with two external electrons; lithium three, and so on. The atomic number is therefore the most important quantity in determining the properties of the elements. When these are arranged in order of atomic number instead of atomic weight, the anomalies of the periodic table disappear. If, then, the atomic number of an element be Z, the electric charge of the nucleus is $+Ze$, and it is surrounded by Z electrons. From Rutherford's experiments on scattering of α rays, it follows that the radius of the nucleus of the hydrogen atom is extremely small in comparison with the orbit of the electron.

It is possible to form an estimate of the size of the electron and of the hydrogen nucleus if the mass of these is considered to be of electromagnetic origin (p. 545). Taking $\mu=1$, mass$=\dfrac{2e^2}{3a}$, and from Millikan's values for the electron (p. 484), $e=1{\cdot}59\times10^{-20}$, and $m=8{\cdot}8\times10^{-28}$,

$$a=\frac{2e^2}{3m}=\frac{2\times1{\cdot}59^2\times10^{-40}}{3\times8{\cdot}8\times10^{-28}}$$
$$=1{\cdot}9\times10^{-13}\text{ cm.}$$

Since the mass of the hydrogen nucleus is 1835 times that of the electron, its radius would be $1{\cdot}0\times10^{-16}$.

From the experiment on scattering by Rutherford it follows that when α particles of range 7 cm. approach to within a distance of $2{\cdot}4\times10^{-13}$ cm. of the hydrogen nucleus, swift hydrogen atoms result.

Hall Effect.—It was found by Hall[2] that when a magnetic field is applied at right angles to a conductor in which an electric current is flowing, a transverse electromotive force arises at right angles to both current and field, so that there is a difference of potential between the two edges of the conductor. In Fig. 446, if the current is parallel to the X axis and the field parallel to the

[1] Van der Broek, *Phys. Zeit.*, xiv., p. 32. 1913.
[2] E. H. Hall, *Phil. Mag.*, **12**, p. 157. 1880.

Y axis, the faces A and B of the conductor are at different potentials. A sheet of the material is employed, the current passing through it from one end to the other, with the magnetic field at right angles to the plane of the sheet. Hall attached terminals to the sides of the sheet, and observed the current in a galvanometer connected to them; but a better method of measuring the effect is to compensate the electromotive force so that no current flows in the galvanometer circuit—in other words, to employ the potentiometer method.

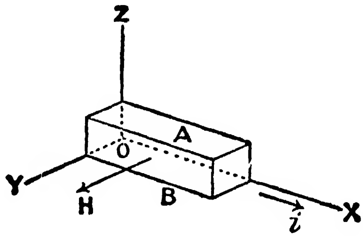


Fig. 446.

The magnitude of this transverse electromotive force is strictly proportional to the current density and to the transverse width of the conductor, and though generally proportional to the strength of magnetic field, this is not always rigidly so.

Thus, difference of potential $= \mathrm{RH}idz$,

where H is the strength of field, i the current density, and dz the width of the conductor at right angles to field and current, R being the coefficient at the Hall effect. The following values are given by Baedeker: [1]—

Gold . . .	$-0{\cdot}00070$	Bismuth . .	-11 (crystallographic axis perpendicular to H)
Silver . .	$-0{\cdot}00088$		
Copper . .	$-0{\cdot}00054$	Bismuth . .	$-0{\cdot}4$ (crystallographic axis parallel to H)
Platinum. .	$-0{\cdot}0002$		
Carbon . .	$-0{\cdot}17$	Antimony .	$+0{\cdot}1$ to $0{\cdot}2$
Iridium . .	$+0{\cdot}00040$	Tellurium .	$+500$ (about)
Zinc . . .	$+0{\cdot}0006$ (about)		

For the ferromagnetic metals, the Hall effect varies in a similar manner to the permeability, disappearing at the critical temperature. For low temperatures, in the case of iron and nickel R is about $+0{\cdot}01$ and for cobalt about $0{\cdot}004$.

With the relative directions of current and field in Fig. 446, B is at a higher potential than A when the Hall effect is positive, and *vice versâ*.

The explanation of the Hall effect by the electronic theory is not entirely satisfactory. With a current flowing in the direction indicated in Fig. 446 the electrons are travelling in the direction XO, and the effect of the magnetic field would be to deflect them downwards. Hence B would on this account be at a lower

[1] K. Baedeker, "Die Elektrischen Erscheinungen in metallischen Leitern."

potential than A, the direction of the transverse electromotive force being that observed in the metals having a negative coefficient of the Hall effect. The occurrence of a positive coefficient in some cases is not easy to account for if the current is carried entirely by electrons having negative charges. If the current were partly due to the motion of positively charged carriers the matter would be easy, for whatever the sign of the carriers the displacement would be downwards in the figure. But in the case of positive carriers, this would raise the potential of B above that of A, and the sign of the Hall effect would be that observed in antimony, etc. We can only say that the electronic theory is at present too incomplete to account properly for all the phenomena of conduction.

Nernst and Ettinghausen Effect.—On maintaining a temperature gradient in a metal sheet, it was found by Nernst and Ettinghausen [1] that in presence of a transverse magnetic field, applied as in Hall's experiment, a potential difference exists between the edges of the sheet. This might be expected upon the electronic theory, for both heat and electrical conduction are supposed to be due to transmission of the electrons. Those moving from the hotter part of the metal have greater velocity, and are therefore more deflected by the magnetic field, than the more slowly moving electrons from the cooler parts.

$$\text{Difference of potential} = Q \cdot H \cdot \frac{dT}{dx} dz,$$

where Q is the coefficient of the Nernst and Ettinghausen effect.

Ettinghausen Effect.[2]—A temperature difference is established between the edges of the plate along which a current is flowing, when there is a magnetic field at right angles to the plane of the plate. This effect is much smaller than the last described, and might have been expected from the transverse deflection of the electrons which gives rise to the Hall effect.

$$\delta T = -\frac{P \cdot Hi}{dy} dz.$$

Leduc Effect.[3]—In this case there is a transverse difference of temperature between the edges of the plate when a temperature gradient exists in the magnetic field.

$$\delta T = SH \cdot \frac{dT}{dz} \cdot dz.$$

[1] H. W. Nernst and A. von Ettinghausen, *Wied. Ann.*, **29**, p. 343. 1886.
[2] A. von Ettinghausen, *Wied. Ann.*, **31**, p. 737. 1887.
[3] A. Leduc, *Journ. d. Phys.*, **6**, p. 378. 1887.

The above four effects have been investigated by Zahn,[1] who gives the following values for the four respective coefficients :—

	R.	S.	Q.	P.
Platinum .	$-1{\cdot}27\times10^{-4}$	$-2{\cdot}1\times10^{-8}$	Very small	Not measurable
Copper .	$-4{\cdot}28\times10^{-4}$	$-23{\cdot}2\times10^{-8}$	$+2{\cdot}7\times10^{-4}$	
Silver . .	$-8{\cdot}97\times10^{-4}$	$-40{\cdot}4\times10^{-8}$	$+4{\cdot}3\times10^{-4}$	
Zinc . .	$+10{\cdot}4\times10^{-4}$	$+12{\cdot}9\times10^{-8}$	$+2{\cdot}4\times10^{-4}$	
Iron . .	$+10{\cdot}8\times10^{-4}$	$+39\times10^{-8}$	$+10{\cdot}5\times10^{-4}$	$-5{\cdot}7\times10^{-8}$
Steel . .	$+133{\cdot}6\times10^{-4}$	$+68{\cdot}7\times10^{-8}$	$+16{\cdot}6\times10^{-4}$	$-6{\cdot}7\times10^{-8}$
Nickel I .	$-46{\cdot}9\times10^{-4}$	-20×10^{-8}	-13×10^{-4}	$+2{\cdot}8\times10^{-8}$
,, II .	-125×10^{-4}	-55×10^{-8}	$-35{\cdot}5\times10^{-4}$	$+17{\cdot}6\times10^{-8}$
Antimony.	$+2190\times10^{-4}$	$+202\times10^{-8}$	-176×10^{-4}	$+134\times10^{-8}$

Longitudinal Effects.—It would follow from the consideration of the above-described transverse effects that the electrons will, due to the transverse motion imposed upon them, now experience similar forces still at right angles to the magnetic field and to their present motion, which therefore makes the effect parallel to the original current or temperature gradient. Thus in the Ettinghausen and the Leduc effects, the upper edge of the plate is at a different temperature to the lower, the electrons travelling from the hotter part have greater velocity and are deflected more by the magnetic field than those from the colder part, thus giving rise to a longitudinal temperature effect. Also in the case of the Hall effect, the lateral displacement of the electrons causes them to experience a force due to the magnetic field, which force is in this case parallel to the original current. The motion of the electrons will thus resemble the motion of the ions in a magnetic and electric field at right angles to each other (p. 474) ; their path is curved, the component of their velocity in the direction of the electric field being less the stronger the transverse magnetic field. The result is therefore to reduce the current, or, in other words to increase the resistance of the conductor. This increase in resistance has been observed in several cases, particularly in that of bismuth, in which metal the Hall effect is very great.

The increase in resistance is independent of the direction of the magnetic field, and may therefore be taken as proportional to the square of the field strength. In the relation—

$$\frac{\delta r}{r}=AH^2,$$

the constant A has been determined for a number of materials and is of the order 10^{-12}, varying from $0{\cdot}06\times10^{-12}$ in the case of platinum to $2{\cdot}8\times10^{-12}$ for cadmium.

[1] H. Zahn, *Ann. der Phys.*, (3) **14**, p. 886 (1904) ; and **16**, p 148 (1905).

In the case of bismuth the effect is complicated, but is sufficiently great to afford a means of measuring magnetic field by determining the resistance of a standardized conductor of bismuth when situated in the field.

The properties of bismuth are in many ways peculiar, as, for example, the variation in resistance to continuous and to alternating current, discovered by Lenard. The resistance to the alternating current depends on the frequency, being less the higher the frequency. Pallme-König,[1] on examining a bismuth wire in currents of very short duration, found that the resistance for a rising current is always less than that for a falling current

TABLE OF TRANSVERSE AND LONGITUDINAL EFFECTS DUE TO A TRANSVERSE MAGNETIC FIELD.

	Transverse.	Longitudinal.
Electro-electric .	HALL. Transverse electromotive force when current flows at right angles to magnetic field.	Change in electrical conductivity.
Electro-thermal	ETTINGHAUSEN. Transverse difference of temperature when current flows at right angles to magnetic field.	Longitudinal difference of temperature.
Thermo-electrical	NERNST and ETTINGHAUSEN. Transverse electromotive force when flow of heat takes place at right angles to magnetic field.	Longitudinal electromotive force.
Thermo-thermal	LEDUC. Transverse difference of temperature when flow of heat takes place at right angles to magnetic field.	Change in thermal conductivity.

Bohr's Theory of the Atom.—One of the most important steps made in the atomic theory was made by Bohr [2] by applying the quantum theory to the problem of the radiation by the hydrogen atom. A simple atom such as that of hydrogen, consisting of a nucleus and one electron, might be in equilibrium with the electron rotating around the atom with such a velocity that the centrifugal force is equal to the attraction between the nucleus and the electron. In fact, the two would rotate about their centre of gravity as a planetary system. It was at first considered that the motion of the electron would give rise to electromagnetic waves of frequency equal to that of the orbital motion of the electron, but the difficulty then arises, that a radiation means loss of energy by the rotating system, with consequent

[1] Pallme-König, *Ann. d. Phys.*, **25**, p. 921. 1908.
[2] N. Bohr, *Phil. Mag.*, **26**, p. 1. 1913.

decrease in velocity, and the electron would continually approach the nucleus. To meet this difficulty Bohr considered that the radiation takes place according to Planck's quantum theory,[1] which states that whenever there is an interchange of energy between radiation and matter, the radiation or absorption takes place in quanta, which are whole multiples of the quantity $h\nu$, where ν is the frequency of the radiation and h is a universal constant, known as Planck's constant, the value of which is $6{\cdot}55 \times 10^{-27}$. Thus, if E is a simple quantum of energy radiated,

$$E = h\nu,$$

or,

$$h = \frac{E}{\nu}.$$

Thus the dimensions of h are given by

$$[h] = [ML^2T^{-2} \div T^{-1}]$$
$$= [ML^2T^{-1}].$$

Planck's constant may therefore be considered as energy multiplied by time, or as an angular momentum.

If we consider the orbit of the electron to be circular, the centre of gravity of the atom being at the centre of the nucleus, the centrifugal force is $m\omega^2 r$, and the force between electron and nucleus is $\frac{e^2}{r^2}$, where r is the distance of the electron from the centre of the atom, and ω the angular velocity of revolution.

$$\therefore\ m\omega^2 r = \frac{e^2}{r^2},$$

or,

$$m\omega^2 r^3 = e^2 \quad \ldots\ldots \quad \text{(i)}$$

Further, the kinetic energy of the electron is $\frac{1}{2}m\omega^2 r^2$, and the total energy is compounded of this and the potential energy. Considering the electrical potential to be zero at infinite distance from the charge $+e$ of the nucleus, the potential at distance r is $+\frac{e}{r}$, and the potential energy on account of the charge $-e$ is $-\frac{e^2}{r}$. The total energy is therefore

$$W = \frac{1}{2}m\omega^2 r^2 - \frac{e^2}{r},$$

But from (i)

$$m\omega^2 r^2 = \frac{e^2}{r},$$

$$\therefore\ W = \frac{e^2}{2r} - \frac{e^2}{r}$$

$$= -\frac{e^2}{2r} = -\frac{1}{2}m\omega^2 r^2 \quad \ldots\ldots \quad \text{(ii)}$$

[1] M. Planck, *Ann. der Physik*, 4, p. 553. 1901.

If the orbit changes from radius r_1 to radius r_2 the change in energy of the atom is $\frac{e^2}{2r_2}-\frac{e^2}{2r_1}$, and if this is accompanied by radiation of frequency ν, then, according to the quantum theory,

$$\frac{e^2}{2r_2}-\frac{e^2}{2r_1}=h\nu \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

So far the theory fails to account for the fact that the atoms of hydrogen do not radiate light of all possible frequencies, for ν might still have any possible value, by giving an infinite variety of values to r. To meet this difficulty Bohr applied the quantum theory to the motion of the electron itself. It is considered that only certain states, called stationary states, are possible, and that when the atom is in a stationary state there is no radiation ; but in passing from one stationary state to another, radiation occurs, which is given by (iii). It has been seen that h is of the dimensions of an angular momentum, and the assumption is made that the angular momentum of the electron can only change by quanta $\frac{h}{2\pi}$. The foundation for this may be found in the quantum specification of W. Wilson,[1] that $\int p dq=nh$, where q is any ordinate defining the system, p the corresponding momentum, and n a whole number, the integration being taken over a complete period. In the case of the hydrogen atom, with an electron having a circular orbit, we may express the motion in polar coordinates, ϕ, r ; r is a constant, and $\omega=\frac{d\phi}{dt}$, which is also constant. Thus, $q=\phi$, and $p=m\omega r^2$,

$$\therefore \int_0^{2\pi} m\omega r^2 d\phi=nh,$$

or,

$$2\pi m\omega r^2=nh,$$

$$m\omega r^2=\frac{nh}{2\pi}.$$

The whole angular momentum $m\omega r^2$ is therefore an integral number of times $\frac{h}{2\pi}$.

Now from (i), $m\omega^2 r^3=e^2$,

$$\therefore r=\frac{n^2h^2}{4\pi^2e^2m} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv)}$$

and the energy in the stationary state represented by (ii) is

$$\frac{2\pi^2e^4m}{n^2h^2}$$

[1] W. Wilson, *Phil. Mag.*, **31**, p. 156. 1916.

If n_1 and n_2 are two integers, then from (iii)

$$\nu h=\frac{2\pi^2e^4m}{n_1^2h^2}-\frac{2\pi^2e^4m}{n_2^2h^2}$$

or,

$$\nu=\frac{2\pi^2e^4m}{h^3}\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right) \quad . \quad . \quad . \quad . \quad . \quad (v)$$

The justification for the assumptions made lies in the remarkable agreement between the frequencies of radiation given by (v) and the known spectrum of hydrogen. It had long been known that the Balmer series of lines in the visible spectrum of hydrogen could be represented by an equation $R\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right)$, where $n_1=2$ and n_2 is made equal to the successive integers greater than 2, each integer corresponding to a line in the spectrum. From measurements on the spectrum, R, which is known as Rydberg's constant, is found to have the value 109677·691 waves per centimetre. On substituting the values $e=4{\cdot}77\times10^{-10}$, $m=8{\cdot}8\times10^{-28}$, and $h=6{\cdot}55\times10^{-27}$,

$$\frac{2\pi^2e^4m}{h^3}=3{\cdot}21\times10^{15}$$

or, dividing by the velocity of light to obtain corresponding units, the value for R is $1{\cdot}07\times10^5$.

It should be noted that in considering the atom as a planetary system a correction should be made for the fact that the centre of rotation has been considered to be the centre of the nucleus instead of the centre of gravity of the system. A correction for this fact would improve the agreement with the known value of Rydberg's constant.

In the above equation (iv) substitution of the known values of h, e and m gives $r=0{\cdot}53\times10^{-8}$ cm. when $n=1$, this being the most probable value of n for the normal state of the atom. The agreement with the value $1{\cdot}1\times10^{-8}$ for the diameter of the hydrogen molecule as derived from the kinetic theory of gases is interesting.

A further advantage is added to Bohr's theory of the atom by observing that if $n_1=1$ and $n_2=2$, 3, 4 etc., a series is obtained whose frequencies correspond to the Lyman series in the hydrogen ultra-violet spectrum, while the numbers $n_1=3$ and $n_2=4$, 5, 6 etc., give the Paschen series in the infra-red. Later, the spectrum $n_1=4$, $n_2=5$, 6, 7 . . was found by Brackett, and the spectrum $n_1=5$, $n_2=6$, 7, 8 . . . by Pfund. Both these spectra are in the infra-red.

In the case of helium the nuclear charge is 2 and the number of external electrons 2. There are now two possibilities of radiation, for the ionisation may consist first in removing one electron

and then the other. In the latter case the process resembles that for the hydrogen atom, but the nuclear charge is 2 instead of 1, and equation (v) becomes

$$\nu=\frac{8\pi^2e^4m}{h^3}\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right).$$

Such a series is known, which again confirms the theory. The spectra of the more complicated atoms require much more complex analysis.

Application to X-rays.—It was shown by Moseley (p. 497) that the square roots of the frequencies of the characteristic X-rays are related in a linear manner to the atomic numbers of the substances emitting the rays. Now relation (v), p. 597, may be written—

$$\nu=\frac{2\pi^2e^2Z^2e^2m}{h^3}\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right)$$
$$=Z^2R\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right)$$

where Ze is the positive charge on the nucleus of the atom.

Taking the simplest possible values, $n_1=1$ and $n_2=2$,

$$\nu=Z^2R\,.\,\tfrac{3}{4}$$

or,

$$Z=Q\sqrt{\nu}.$$

It is clear from Moseley's curves (p. 497) that if we take the K radiations Z increases by regular amounts as we pass up the series of elements, and it was this fact that led him to the conclusion that the nuclear charge determines the characteristic radiation.

If the case of copper be chosen, then for the $K\alpha_1$ line

$$\lambda=1{\cdot}54\times10^{-8},$$
$$\therefore\ \nu=\frac{3\times10^{10}}{1{\cdot}54\times10^{-8}}=1{\cdot}95\times10^{18}.$$

Also Rydberg's constant $R=1{\cdot}097\times10^5\times3\times10^{10}$,

Z for copper$=29$, and $e=1{\cdot}59\times10^{-20}$,

which give,

$$\nu=29^2\times1{\cdot}097\times3\times10^{15}\times\tfrac{3}{4}$$
$$=2{\cdot}07\times10^{18}.$$

This agreement is surprisingly good, considering the assumptions involved. Moseley showed that there is still a correction to be made for the effect of the electrons upon each other. There is little doubt that the K radiation originates in the innermost ring of electrons, which is called the K ring. Bearing in mind the origin of the radiations, it appears that when an electron moving sufficiently rapidly through an atom removes one of the

electrons from the K ring, the return of an electron from outside, probably the next ring, gives rise to the characteristic X-radiation. The K radiations thus correspond, in the case of the heavier elements, to the Lyman series in the hydrogen spectrum.

Following the analogy, if we put $n_1=2$ and $n_2=3$,

$$\nu=Z^2R(\tfrac{1}{4}-\tfrac{1}{9})=Z^2R\,.\,\tfrac{5}{36}.$$

Taking platinum, $\lambda=1{\cdot}32\times10^{-8}$, and $Z=78$,

$$\nu=\frac{3\times10^{10}}{1{\cdot}32\times10^{-8}}=2{\cdot}27\times10^{18}.$$

and

$$\nu=78^2\times1{\cdot}097\times3\times10^{15}\times\tfrac{5}{36}$$
$$=2{\cdot}78\times10^{18}.$$

It is therefore reasonable to recognise the analogy between the L radiation and the Balmer series of the hydrogen spectrum.

Again, putting $n_1=3$ and $n_2=4$,

$$\nu=Z^2R(\tfrac{1}{9}-\tfrac{1}{16})=Z^2R\tfrac{7}{144}$$

Taking M_α for thorium as $4{\cdot}14\times10^{-8}$

$$\nu=\frac{3\times10^{10}}{4{\cdot}14\times10^{-8}}=0{\cdot}725\times10^{18},$$

and, since $Z=90$,

$$\nu=90^2\times1{\cdot}097\times3\times10^{15}\times\tfrac{7}{144}$$
$$=1{\cdot}24\times10^{18}.$$

It must be remembered that with high atomic numbers the correction for the interaction of the electrons in the rings is very great. Apart from this, the agreement is sufficiently near to recognise the analogy between the M radiations and the Paschen series of the hydrogen spectrum.

From equation (i) (p. 595), it follows with increasing nuclear charge, that e^2 must be replaced by Ze^2 and that $m\omega^2r$ increases. It follows that as r diminishes with increasing attraction, ω, the angular velocity which determines the frequency of revolution of the electrons in their orbits, increases.

Ionisation Potential.—The numbers $n_1=1$ and $n_2=2$ are the smallest integers that can give any meaning to equation (v), and would probably correspond to some limiting condition of energy of the atom. If $n=1$, the energy of the atom is $\frac{2\pi^2e^4m}{h^2}$, and this is generally considered to be the energy in the normal condition. It will also be remembered that the condition of zero energy was taken to be that in which the electron is removed to an infinite distance from the nucleus and is at rest with respect to it, which condition is recognised as that in which the atom is ionised. The

energy $\dfrac{2\pi^2e^4m}{h^2}$ may then represent the work necessary to ionise the atom. On putting in the values $e=4{\cdot}774\times10^{-10}$, $\dfrac{e}{m}=5{\cdot}31\times10^{17}$, and $h=6{\cdot}55\times10^{-27}$, the energy for ionisation is $1{\cdot}703\times10^{-11}$ ergs. This is frequently represented in terms of the potential through which an electron would have to fall in order to acquire this energy.

Now $Ve=\frac{1}{2}mv^2$ ergs, where v is the velocity acquired by the electron in passing through a difference of potential V. The value of e in the electromagnetic units is $1{\cdot}59\times10^{-20}$, and if V is in volts,

$$V\times10^8\times1{\cdot}59\times10^{-20}=1{\cdot}703\times10^{-11},$$
$$V=10{\cdot}8 \text{ volts.}$$

The agreement with the known value is sufficiently good to indicate the validity of the assumptions made.

Ionisation potential takes a special importance from the relation between X-rays and the electrons producing them and the electrons produced by X-rays. For the production of any given quality of X-ray, the velocity of the cathode ray in the X-ray tube must have a certain minimum value. This minimum velocity in the case of the K radiation was found by Whiddington (p. 488) to be $2(Z-2)10^8$ cm. per sec., where Z is the atomic number of the substance of which the anticathode is constructed. In the case of nickel, $Z=28$, so that 52×10^8 is the velocity of electron required to produce the K radiation. The energy of the electron is therefore $\frac{1}{2}mv^2=\frac{1}{2}\times8{\cdot}8\times10^{-28}\times(52\times10^8)^2=1{\cdot}2\times10^{-8}$ ergs. If this energy is converted into radiation and the quantum theory applies to the process,

$$1{\cdot}2\times10^{-8}=h\nu$$
$$\therefore\ \nu=\frac{1{\cdot}2\times10^{-8}}{6{\cdot}55\times10^{-27}}$$
$$\text{and } \lambda=\frac{3\times10^{10}}{\nu}=\frac{6{\cdot}55\times10^{-27}\times3\times10^{10}}{1{\cdot}2\times10^{-8}}$$
$$=1{\cdot}64\times10^{-8}.$$

The observed values of λ for the two K radiations of nickel are $1{\cdot}66\times10^{-8}$ and $1{\cdot}51\times10^{-8}$, so that the agreement is sufficiently close to justify the application of the quantum theory to this exchange of energy. Whiddington [1] has also employed the method of Rutherford and Robinson (p. 521) to the examination of the velocity of the electrons emitted by various substances when subjected to X-rays. He finds that the fastest moving

[1] R. Whiddington, *Phil. Mag*. 43. p. 1116. 1922.

electrons have velocities very well in accordance with the frequency of the X-rays producing them, as determined by the relation $\frac{1}{2}mv^2=h\nu$.

Energy Levels.—There is a very convenient method of representing spectra, which exhibits all the lines indicated by the simple Bohr theory for a given atom. For example, in the case of the hydrogen atom equation (v) (p. 597) gives the frequencies for the lines by taking integral values of n_1 and n_2. From the line AB (Fig. 447) let the distance AC be taken to scale to represent $\frac{2\pi^2e^4m}{h^2}$, and AE to be $\frac{2\pi^2e^2m}{2^2h^2}$ and AG to be $\frac{2\pi^2e^2m}{3^2h^2}$.

Then the distance CE represents $\frac{2\pi^2e^4m}{h^2}\left(\frac{1}{1^2}-\frac{1}{2^2}\right)$, which is the energy emitted by an atom that gives rise to the first Lyman line. Similarly GC represents the energy $\frac{2\pi^2e^4m}{h^2}\left(\frac{1}{1^2}-\frac{1}{3^2}\right)$. EF, GH, etc., are called energy levels and from them the spectrum can be constructed. Thus if the electron in the hydrogen atom in the normal condition absorbs an amount of energy

$$\frac{2\pi^2e^4m}{h^2}\left(\frac{1}{1^2}-\frac{1}{4^2}\right)$$

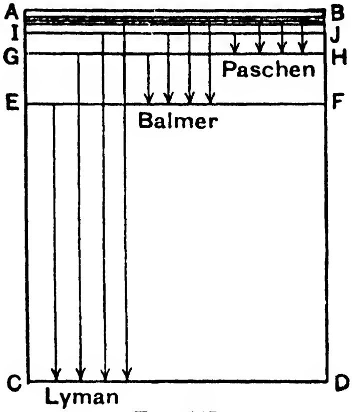


FIG. 447.

it is raised from the zero energy level CD to the energy level IJ. It may now return to the original condition in a variety of ways. In doing it in one step, energy CI$=\frac{2\pi^2e^4m}{h^2}\left(\frac{1}{1^2}-\frac{1}{4^2}\right)$ is liberated and the third Lyman line is emitted. Or it may fall in two steps IE and EC, giving the second Balmer line and the first Lyman line simultaneously. Or, again it may fall from the level IJ to level GH and then alternatively by step GC or steps GE and EC. Thus the first Paschen and second Lyman line may be emitted, or the first Paschen, first Balmer and first Lyman lines.

For atoms more complex than that of hydrogen, diagrams of energy level may be drawn, but they are more complex than that for the hydrogen atom.

The differences in energy level may be expressed in electron-volts (p. 510). Thus for the first Lyman line the wave-length is $1{\cdot}220\times10^{-5}$ cm. Thus $\nu=\frac{3\times10^{10}}{1{\cdot}220\times10^{-5}}$, $h\nu=\frac{6{\cdot}55\times10^{-27}\times3\times10^{10}}{1{\cdot}220\times10^{-5}}$

ergs. An electron of charge $1{\cdot}59\times10^{-20}$ falling through V volts would acquire $1{\cdot}59\times10^{-20}\times10^{8}$ V ergs. Equating these, V=10·13. Thus CE=10·13 electron-volts which fixes the scale for the diagram. Note that 1 electron-volt$=1{\cdot}59\times10^{-12}$ ergs.

The Bohr Magneton.—On p. 596 it was seen that the angular momentum of the electron consists in multiples of the quantity $\frac{h}{2\pi}$. Now the angular momentum is $m\omega r^2$ and the magnetic moment of a circular orbit is $\frac{\pi r^2 e}{T}$ (p. 547). The ratio of the magnetic moment to the moment of momentum is therefore $\frac{\pi r^2 e}{Tm\omega r^2}=\frac{\pi e}{Tm\omega}$. But $T=\frac{2\pi}{\omega}$, so that the ratio is $\frac{e}{2m}$. Since the angular momentum consists of multiples of $\frac{h}{2\pi}$, the magnetic moment consists of multiples of $\frac{e}{2m}\cdot\frac{h}{2\pi}=\frac{eh}{4\pi m}$, which is known as the *Bohr magneton*. Taking $e/m=1{\cdot}757\times10^{7}$ and $h=6{\cdot}55\times10^{-27}$, the value of the Bohr magneton is $9{\cdot}16\times10^{-21}$. Taking the value for one gramme atom instead of one atom the value is $9{\cdot}16\times10^{-21}\times6{\cdot}064\times10^{23}=5550$. Taking the value of the Weiss magneton as 1125·6 (p. 575) the ratio of the Bohr magneton to the Weiss magneton is 5550/1125·6=4·93.

Magnetic Deviation of Atom Streams.—On p. 236 it was found that in a magnetic field which varies from point to point there is a resultant force on any magnet. Following similar reasoning it appears that if an atom has a magnetic moment it will experience a force when situated in a non-uniform magnetic field. Imagine an atom to have a pole strength p with distance dx between the poles. If the magnetic axis is in the direction of the field H, the force on one pole is pH and on the other $pH+p\cdot\frac{dH}{dx}dx$, so that the resultant force is $\frac{dH}{dx}\cdot pdx=m\frac{dH}{dx}$, where m is the magnetic moment $p\,.\,dx$. The force is directed towards the stronger or the weaker parts of the field according to the directions of field and magnetic moment. A fine beam of atoms passing through such a magnetic field would then be split up in a manner determined by the distribution in direction of the magnetic axes of the atoms. If the distribution is in all directions the beam would be broadened, but if, as is required by the quantum theory, the magnetic axes can only make certain angles with the field, the beam is split into parts one corresponding to each such angle. Stern and Gerlach [1] heated the metal (in the first

[1] O Stern and W. Gerlach. *Ann. der Phys.*, 76. p. 163. 1924.

case silver) in a small oven and allowed the stream issuing from a small aperture to be further limited by slits and then to pass through a non-uniform magnetic field. The field is produced by an electromagnet with pole pieces, one of which has a sharp edge opposite and parallel to a groove in the other. The beam after passing through the field, falls upon a glass or metal target and condenses there giving a trace which may be developed if necessary. The experiments showed a definite splitting of the beam into two, one on each side of the zero field position, thus definitely establishing the validity of the quantum theory. The attracted beam is more diffuse than the repelled beam, owing to a variation of $\frac{dH}{dx}$ near the sharp edge of the pole piece, and the deflection of the repelled beam is used for purposes of calculation. From the relations $F=m\frac{dH}{dx}$, and $s=\frac{1}{2}\frac{F}{\text{mass}}\left(\frac{l}{v}\right)^2$, where l is the length of track in the field, s the displacement of the beam and v the velocity of the particles, m the magnetic moment is found. For silver the value of the gramme atom magneton was found to be 5690, which agrees fairly well with that of the Bohr magneton (p. 602).

Photoelectricity.—The Hallwachs phenomenon (p. 473) is only one case of the liberation of electrons when light falls upon matter. The effect is very widely observed, and to it is applied the term *photoelectricity*. Experiments conducted in a vacuum have given us much more intimate knowledge of the process than we had previously, and have led to the discoveries that (*a*) the velocity of the electrons emitted is independent of the intensity of the light, and (*b*) the rate of emission of the electrons is directly proportional to the intensity of the light. A measure of the positive potential acquired by the illuminated plate enables the velocity of the emitted electrons to be found, for the emission ceases when the electric intensity produced by the loss of negative electricity is sufficient to prevent the further escape of electrons. Also the saturation current (p. 477) for a given p.d. between the illuminated plate and a parallel plate gives the number of electrons emitted per second. In the case of the alkali metals maximum photoelectric effect occurs for light belonging to the visible part of the spectrum. This is probably due to a *selective* effect for the metal, as, for the *normal* photoelectric effect, the shorter the wave-length of the light the greater is the emission. The plane of polarisation of the incident light also influences the rate of emission of the corpuscles.

The photoelectric phenomenon has been shown to be connected with those of fluorescence and phosphorescence as well as with that of the chemical changes occurring in the photographic plate.

The subject is now such a large and important one that the student can only be referred to such works on it as that of H. S. Allen.[1]

The above condition (*a*) indicates that the velocity of emission of electrons depends only upon the frequency of the incident light, and many observers concur in the conclusion that it is independent of the intensity of the light and the temperature of the metal. For any given metal there is a minimum frequency for the incident light, below which there is no emission of electrons. The electrons emitted have various velocities up to a certain maximum, but as those emitted below the surface probably lose energy in passing through the material, it is reasonable to suppose that the velocity with which the electrons are liberated from the atoms is the same for them all, and equal to the maximum observed velocity v. If V is the potential through which the electron must fall to acquire its velocity of emission, the equation $Ve=k\nu-w_0$, has been found to fit the relation between $Ve(=\frac{1}{2}mv^2)$ and ν the frequency of the incident light. w_0 has a definite value for each metal, and k is a constant, having the same value for all metals. The quantum theory has been applied by Einstein [2] to the photoelectric emission, by considering k to be identical with Planck's constant h. The atom can then absorb light only in complete quanta of energy ($h\nu$), and if ν_0 is the minimum frequency of light which will produce liberation of electrons, $h\nu_0$ is the energy absorbed from the incident beam to liberate electrons with zero velocity. This is identified as the quantity w_0, and the equation becomes—

$$Ve=\tfrac{1}{2}mv^2=h(\nu-\nu_0).$$

Experiments by Richardson [3] have proved the validity of this equation, and Millikan [4] has employed it for the calculation of h. It thus appears that when a light quantum $h\nu$ is absorbed, the energy $h\nu_0$ is used in liberating the electron and the balance of energy $h\nu-h\nu_0$ is available for producing kinetic energy, which may be distributed between the atom and the electron.

The phenomenon of ionisation by X-rays may be looked upon as an example of the photoelectric effect, for it is now known that X-rays are of the same nature as light, but of higher frequency. This explains the greater ionisation produced by rays of higher frequency; for the quantum of energy absorbed by the atom is $h\nu$, and is therefore proportional to the frequency. It also explains the fact that the corpuscular rays emitted when

[1] H. S. Allen, "Photo-electricity." Longmans, Green & Co.
[2] A. Einstein, *Ann. d. Physik*, **17**, p. 132. 1905.
[3] O. W. Richardson, *Phil. Mag.*, **24**, p. 575. 1912.
[4] R. A. Millikan, *Phys. Rev.*, vii, **18**, p. 355. 1916.

X-rays pass through gases, have the same velocity as the electrons which gave rise to the X-rays (p. 471).

Photoelectric Cells.—There are several forms of photoelectric cell. A typical *photo-emissive cell* has in its glass envelope a curved cathode A (Fig. 448), often of hemi-cylindrical form, and a wire anode B. The battery provides a field to ensure that electrons released from A all reach B. The small resultant current flows through the high resistance CD across which a potential difference appears, which may be amplified by thermionic valves (p. 585).

The material of the cathode determines the sensitivity of the device and the range of wave-lengths which affect it. Zinc responds to ultra-violet; the alkali metals are much more sensitive and respond also to visible radiation. A suitable discharge in hydrogen was shown by Elster and Geitel[1] greatly to enhance

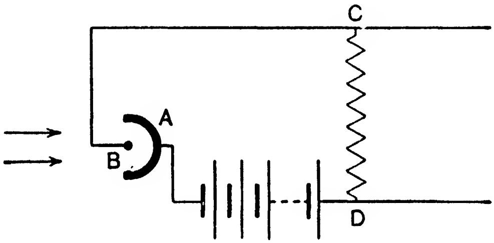


FIG. 448.

the sensitivity. Thin layers of alkali metals or their oxides on suitable bases give even better results, especially if the layers are only one atom thick. The cæsium cell, due to Koller,[2] has maximum response for wave-length 0·000055 cm., in the middle of the visible spectrum.

An evacuated cell gives a linear response, the saturation current being proportional to the intensity of the incident radiation. With a gas such as argon present, ionisation by collision (p. 506) increases the current, but the resultant current is no longer proportional to the illumination.

Photo-emissive cells are put to many uses, among which may be mentioned their employment to obtain currents fluctuating to correspond with the light passed by the sound-track on the margin of a cinematograph film, the electrical signals then being amplified and applied to loud-speakers which emit the sound.

A *photo-conductive cell* utilises the change in resistance of a semiconductor on irradiation, produced by electrons liberated in the interior of the substance, an internal photoelectric effect. Selenium, at one time extensively used in this way, is sluggish

[1] J. Elster and H. Geitel, *Phys. Zeit.*, **14**, p. 741. 1913.
[2] L. R. Koller, *Jn. Opt. Soc. Amer. and Rev. Sci. Instr.*, **19**, p. 135. 1929.

in response. Nowadays thin layers of suitable materials are used: for example lead sulphide, which responds well to infra-red radiation. The sensitive element may be used in place of the cell AB in a circuit similar to that of Fig. 448.

The *barrier-layer cell* is a third type, a typical example of which has a copper base heavily oxidised to cuprous oxide (Cu_2O), this being in turn coated with a very thin layer of copper, deposited by evaporation in a vacuum from an electrically heated piece of copper. Photo-electrons released from this copper layer mostly pass to a copper ring round its periphery since the copper/copper oxide interface has a rectifying action (p. 452), offering a higher resistance to electrons passing from copper into the oxide than for the reverse direction. A current will therefore flow in an external circuit connecting the copper electrodes without any battery being necessary. Barrier-layer cells may be connected direct, or through fairly low resistances, to sensitive moving-coil indicators and are so used in photometry, some photographic exposure meters providing examples of this use.

Absorption.—On the dynamical theory of radiation it would be expected that when a continuous beam of light, such as white light, passes through matter, the atoms would absorb just that frequency of radiation which they themselves emit; the reason for this is that by resonance they will be set in oscillation and so absorb energy from the incident beam. But radiation has been accounted for (p. 596) by the passage of the atom from its ionised state to its normal state through a series of stationary states. We should therefore expect that the absorption of energy would transfer the atom back through this process from the normal to the ionised condition. The emission of radiation of a given frequency implies the presence of two stationary states, and absorption of radiation of this frequency will only occur if the one of these states corresponding to lower energy is present in some of the atoms of the substance. Since in the ordinary condition of a substance only the normal unionised condition is likely to exist, the only absorption that will occur is of such a frequency that the normal state is the stationary state of lowest atomic energy. Absorption corresponding to states intermediate between the normal and the ionised condition may occur. Thus the lowest frequency of radiation for absorption is the limiting frequency given by—

$$h\nu = W_1 - W_N$$

where W_1 is the energy of the atom in the ionised state and W_N that in the normal state. If ν is greater than this frequency, each quantum of energy absorbed is in excess of that required for ionisation and photoelectric emission results (p. 604). It does

not follow that the atom once ionised will return in one step to its ionised condition ; in fact, it returns through the stationary states and emits the series defined by equation (v) (p. 597), which correspond to fluorescence. The above reasoning explains why the absorption spectrum in the case of hydrogen does not correspond to the emission spectrum. Since the condition of lowest atomic energy corresponding to (v) is W_1, obtained by putting $n_1=1$, the possible absorption frequencies are obtained by putting $n_1=1$, and $n_2=2, 3, 4$, etc., that is, they correspond to the Lyman emission series. The series $n_2=2$, and $n_3=3, 4$, etc., will not correspond to an absorption spectrum, because atoms corresponding to the stationary state $n_2=2$ are not present in the gas under ordinary conditions.

It is known that a certain critical velocity of cathode ray is necessary for the production of X-rays of any given frequency (p. 600). This velocity is generally given in terms of the voltage through which the electron must pass to give it this velocity.

Since $\frac{1}{2}mv^2=\text{V}e$, and $\dfrac{e}{m}=1{\cdot}76\times10^7$ (p. 484),

$$v^2=2\times1{\cdot}76\times10^{15}\ \text{V}$$

or,
$$v=5{\cdot}92\sqrt{\text{V}}\,10^7 \text{ cm. per sec.}$$

and,
$$\text{V}=2{\cdot}80v\ 10^{-16} \text{ volts.}$$

It is found that on increasing the voltage in the X-ray tube the general emission of X-rays increases, but there is a sharp upper limit to the frequency for any voltage, which limit rises with the voltage for any given substance. Also in the absorption of X-rays by any substance there is a sharp limit to the absorption for the K radiation, known as the K *absorption limit.* On the quantum theory it would be expected that the absorption limit is given by the relation $h\nu=\text{W}_1-\text{W}_\text{N}$, where W_1 is now the work required to remove one electron from the K ring. In returning to the normal condition the greatest frequency of radiation emitted is also given by $h\nu=\text{W}_1-\text{W}_\text{N}$, which is the voltage necessary to produce the rays.

$$\therefore\ \tfrac{1}{2}mv^2=\text{V}e=h\nu.$$

Using the values $e=1{\cdot}59\times10^{-20}$, and $h=6{\cdot}55\times10^{-27}$,

$$\lambda=\frac{3\times10^{10}}{\nu}=\frac{6{\cdot}55\times10^{-27}\times3\times10^{10}}{1{\cdot}59\times10^{-20}\times\text{V}\times10^8}$$
$$=\frac{1{\cdot}24\times10^{-5}}{\text{V}}$$

where λ is in centimetres and V in volts.

Photons and Compton Effect.—It has been seen that whenever atomic radiation of waves occurs, the amount of energy radiated is $h\nu$ or multiples of this quantity. This is the essential feature of the quantum theory; that $h\nu$ is a distinct bundle or corpuscle of energy, now called a *photon,* and that in any process of atomic radiation only complete photons are radiated. According to Einstein's theory of relativity, matter and energy are interchangeable, just as the various kinds of energy are mutually convertible, and that in the conversion between mass and energy the amount of energy equivalent to a mass m is mc^2. Thus if a photon is associated with a mass m, this is given by the relation $mc^2=h\nu$.

Another property of mass is that when in motion it possesses momentum, and that in the collision between two masses the momentum is conserved. In the case of the photon the momentum is $mc=\frac{h\nu}{c}$. If this impinges upon a free electron at rest, the electron gains momentum and the photon loses it. The photon therefore loses energy and the product $h\nu$ is thus diminished. This can only happen by the frequency ν being lowered. This is known as the *Compton effect.*

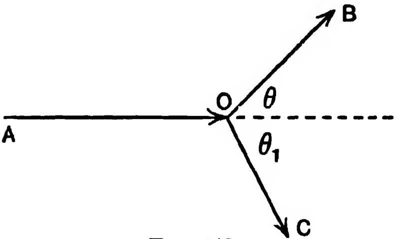


FIG. 449.

Let AO (Fig. 449) represent the momentum $\frac{h\nu_0}{c}$ of the impinging photon, OB the momentum $\frac{h\nu_\theta}{c}$ of the scattered photon and OC that of the electron which was originally at rest at O. If m is now the mass of the electron at rest and v its velocity along OC, its effective mass is $\frac{m}{\sqrt{1-\beta^2}}$, where $\beta=\frac{v}{c}$ (p. 546) and its momentum is $\frac{mv}{\sqrt{1-\beta^2}}=\frac{m\beta c}{\sqrt{1-\beta^2}}$.

For the momentum to be conserved

$$\frac{h\nu_0}{c}=\frac{h\nu_\theta}{c}\cos\theta+\frac{m\beta c}{\sqrt{1-\beta^2}}\cos\theta_1$$

$$\frac{h\nu_0}{c}-\frac{h\nu_\theta}{c}\cos\theta=\frac{m\beta c}{\sqrt{1-\beta^2}}\cos\theta_1$$

and,
$$\frac{h\nu_\theta}{c}\sin\theta=\frac{m\beta c}{\sqrt{1-\beta^2}}\sin\theta_1.$$

Squaring these two equations and adding,

$$\left(\frac{h\nu_0}{c}\right)^2+\left(\frac{h\nu_\theta}{c}\right)^2-2\frac{h\nu_0}{c}\frac{h\nu_\theta}{c}\cos\theta=\frac{m^2\beta^2c^2}{1-\beta^2}.$$

If there is no loss of energy in the process, the original energy of the photon, $h\nu_0$, must equal the sum of the energy of the scattered photon $h\nu_\theta$ and kinetic energy of the electron. This is $mc^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right)$, a quantity which reduces to $\frac{1}{2}mv^2$ when v is small compared with c, for it may be written $mc^2\frac{1-(1-v^2/c^2)^{\frac{1}{2}}}{(1-v^2/c^2)^{\frac{1}{2}}}$, and the numerator expanded gives $\frac{1}{2}\frac{v^2}{c^2}$ when v is small. Another way of looking at the same fact is to remember that the energy of rest mass is mc^2 and since the mass when in motion is $\frac{m}{\sqrt{1-\beta^2}}$, its corresponding energy is $\frac{mc^2}{\sqrt{1-\beta^2}}$. The excess $\frac{mc^2}{\sqrt{1-\beta^2}}-mc^2$ or $mc^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right)$ is thus the kinetic energy.

The energy equation is therefore

$$h\nu_0-h\nu_\theta=mc^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right).$$

Dividing by c and squaring—

$$\left(\frac{h\nu_0}{c}\right)^2+\left(\frac{h\nu_\theta}{c}\right)^2-2\frac{h\nu_0}{c}\cdot\frac{h\nu_\theta}{c}=m^2c^2\left(\frac{1}{1-\beta^2}+1-\frac{2}{\sqrt{1-\beta^2}}\right)$$

and subtracting this from the momentum equation

$$\frac{2h\nu_0}{c}\cdot\frac{h\nu_\theta}{c}(1-\cos\theta)=\frac{m^2c^2}{1-\beta^2}(\beta^2-1)-m^2c^2+\frac{2m^2c^2}{\sqrt{1-\beta^2}}$$

$$=2m^2c^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right)$$

$$2\frac{h\nu_0}{c}\cdot\frac{h\nu_\theta}{c}\,2\sin^2\frac{\theta}{2}=2m(h\nu_0-h\nu_\theta)$$

$$\nu_\theta=\nu_0\bigg/\left(1+2\frac{h\nu_0}{mc^2}\sin^2\frac{\theta}{2}\right).$$

Using the characteristic X-rays of molybdenum Compton [1] allowed them to fall upon graphite, which contains free electrons. Compton and subsequent experimenters have found the change

[1] A. H. Compton. *Phys. Rev.*, xxi, pp. 483, 715 (1923); xxii, p. 409 (1923).

in frequency of X-rays on scattering by free electrons to occur in agreement with the above equation.

Atomic Structure.—From the account that has already been given (p. 594) of the Bohr theory of the atom, it will be realised that the atomic number Z, which defines the electrical charge of the nucleus, determines the number of electrons occupying orbits round the nucleus. All the available evidence suggests that no other property of the nucleus has any major influence on the extra-nuclear electrons. It is therefore convenient to discuss the outer structure first and to make the structure of the nucleus a separate issue.

Each type of atom possesses an array of possible electron orbits which may be grouped into " shells " known as the K, L, M, . . . shells. In the undisturbed neutral atom, only those orbits of lowest energies will be occupied by electrons. Energy may be temporarily stored by the atom if an electron passes to a higher energy level. The absorption and emission of radiation by a hydrogen atom was discussed on p. 601 in this way. Fig. 450 illustrates diagrammatically the scheme for a more complex atom such as an alkali metal.

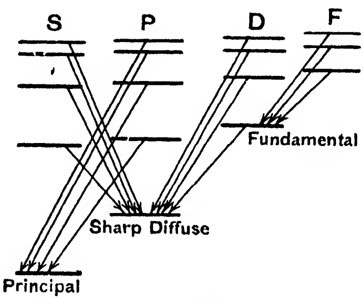


FIG. 450.

The names *sharp*, *principal*, *diffuse* and *fundamental* and their initial letters *s*, *p*, *d*, *f*, originally applied by spectroscopists to distinguish the spectral emission lines in the various series, have become transferred to the sub-divisions of the main energy-levels or shells and to the electrons occupying them. Thus the K shell is complete with only 2 electrons, but the L shell may have up to 8 and these are grouped as 2 *s* and 6 *p* electrons. The M shell may have up to 2 *s*, 6 *p* and 10 *d* electrons, a total of 18; and so on.

Chemical combination of atoms to form molecules is presumed to involve interactions between the outer parts only of the electronic structures of the reacting atoms. The extreme chemical inertness of the group of gases given below implies great stability in their electronic structures.

THE INERT GASES.

Element	He.	Ne.	A.	Kr.	Xe.	Rn.
Atomic number (Z) .	2	10	18	36	54	86
Electronic structure (additional to that preceding)	K2	+L8 (2*s*, 6*p*)	+M8 (2*s*, 6*p*)	+M10 (10*d*) +N8 (2*s*, 6*p*)	+N10 (10*d*) +O8 (2*s*, 6*p*)	+N14 (14*f*) +O8 (2*s*, 6*p*)

In the foregoing table, completion of each successive shell is marked by a horizontal line.

All the elements which have one electron more than one of these inert gases are members of the group of alkali metals (3 Li, 11 Na, 19 K, 37 Rb, 55 Cs, 87 Fr). They are all strongly electro-positive, forming singly-charged ions. This behaviour is explained by the readiness with which the single extra electron is lost, leaving behind a stable structure very similar to that of the preceding inert gas. These elements are in turn followed by the alkaline-earth metals (Be, Mg, Ca, Sr, Ba, Ra), which are electro-positive with valency 2, corresponding to the loss of both extra electrons. Correspondingly, the element with Z one less than for an inert gas is a halogen (F, Cl, Br, I, At), univalent and electro-negative, readily accepting an extra electron to form a more stable configuration. It has already been seen (p. 493) how the atoms of the two elements of an alkali halide such as NaCl alternate in the crystal lattice. These "atoms" are in fact the ions, Na^+ and Cl^-, which also exist in aqueous solution. Such ionic crystals are usual with inorganic salts. Presumably the electrostatic forces between the ions play a dominant part in holding them together.

In other structures the electrons must be allotted more direct parts in the binding of the whole. Thus in a simple diatomic molecule, the atoms may be pictured as having an orbit in common, occupied by a pair of electrons, one from each atom. This simple idea must not be taken too literally, but must suffice at this stage. Metals might be regarded as an extreme case of sharing, with the outer electrons no longer bound to one or to a pair of atoms, but shared by a large group. The typical metallic crystal thus contains regularly placed metallic ions, with their mutual repulsions rendered ineffective by the cloud of electrons moving between them.

Quantum Numbers.—The principle of quantisation of the orbits in the Bohr atom (p. 596) was extended by W. Wilson from circular to elliptical orbits. Two quantum numbers are necessary to specify an elliptical orbit, equivalent to giving both major and minor axes of the ellipse, but according to classical mechanics the energy of the electron in the orbit should depend on the major axis alone. To explain the fine structure revealed by lines of the hydrogen spectrum when examined under high resolving power, Sommerfeld pointed out that the high velocity of an electron when in the part of an elliptical orbit nearest to the nucleus necessitates a variation in its mass, according to the principle of relativity, and this entails a change in the energy associated with the orbit. With other atoms there is an even more marked effect, as an electron in a very eccentric orbit may penetrate the shell beneath its own and come into a much more intense field than when in the rest of its orbit, partially shielded from the nucleus.

Thus electrons belonging to the one shell specified by a *principal quantum number n*, may be further grouped into sub-shells characterised by a second quantum number, nowadays usually described as the *orbital angular momentum number l*. The value $l=0$ gives the *s* electrons, referred to already; similarly $l=1$ gives *p*, $l=2$ corresponds to *d*, $l=3$ to *f*, and so on for *g*, *h*, . . .

Two further quantum numbers are required to complete the specification of the state of an electron in any atom. These may be given in more than one way, but one of these numbers may be allotted with reference to the orientation which an electronic orbit adopts relative to a magnetic field. A circular orbit of radius r, traversed by the charge e some $v/2\pi r$ times per second, is equivalent to a current $ev/2\pi r$ round the orbit. Treating the orbit as a magnetic shell (p. 225), it has a magnetic moment $\mu=\frac{1}{2}evr$, while its angular momentum $p=mvr$, so that the ratio $\mu/p=\frac{1}{2}(e/m)$. This last relationship is true for all shapes of orbit. According to quantum laws, the component of angular momentum parallel to the field must be an integral multiple of $h/2\pi$ and the multiple is specified by the *magnetic quantum number* m_l. Applied to the Zeeman effect (p. 548), these ideas have not only explained the patterns of lines observed in the spectra, but have also yielded a value of e/m for the electron which is in close agreement with other estimates.

The *spin magnetic quantum number* m_s may be selected for the fourth number. This refers to the hypothesis of Uhlenbeck and Goudsmit,[1] advanced in 1925, that the electron has an intrinsic angular momentum and magnetic moment similar to those which a spinning electric charge might be expected to have. Theory assigns the value e/m to the ratio μ/p, twice that for orbital motion, and the permissible steps in angular momentum are one-half those for orbital motion and are thus changes of $\frac{1}{2}(h/2\pi)$. In a magnetic field, the spin aligns itself either parallel to the field or anti-parallel and the spin magnetic quantum number $m_s=\pm\frac{1}{2}$. The parallel arrangement has lower energy than the anti-parallel, just as an ordinary magnet has less energy when aligned with a magnetic field.

The array of four quantum numbers (n, l, m_l, m_s) defines the state of an electron in an atom. Even in the absence of any applied field, the magnetic numbers are effective, possibly because of the field produced by the nucleus. The possible values which the various numbers may have are subject to certain restrictions: for example, l may have integral values from 0 to $n-1$; and certain Selection Rules govern the permissible transitions of an

[1] G. E. Uhlenbeck and S. Goudsmit, *Naturwiss*, **13**, p. 593 (1925); *Nature*, **117**, p. 264 (1926).

electron from one state to another: *e.g.* l changes by ± 1. These rules have helped considerably in reducing to order the great mass of spectroscopic data, but it must be admitted that "forbidden lines," corresponding to transitions not according to simple rules, are by no means unknown.

Pauli's Exclusion Principle.—The simplicity of the scheme of numbers of electrons in each complete shell of an atom invites an explanation and this is provided by Pauli's exclusion principle,[1] according to which no one quantum state, specified by its set of quantum numbers, can be simultaneously occupied, in any one atomic system, by more than one electron.

The spin quantum number doubles all the possibilities afforded by the first three quantum numbers. In pictorial language, each orbit may be occupied by a pair of electrons which have their spins oppositely directed. The principal quantum number n defines the shell ($n=1$ for the K shell, $n=2$ for L, etc.) and it can be shown from the Pauli principle that the limitations imposed on the other numbers permit a maximum number of electrons for the nth shell of $2n^2$.

The Periodic Table.—The recognition of a periodic recurrence of the chemical and physical properties of the elements when they are set out in the order of atomic weights was made by Newlands, Mendeléeff (1869) and others and successful predictions of properties of elements then unknown were made on this basis, especially by Mendeléeff. In addition the strict order of ascending atomic weights was inverted in certain cases (*e.g.* Co, atomic weight 58·9, was placed before Ni, atomic weight 58·7) and subsequent work on atomic number (p. 497) has confirmed these changes. A form of periodic table is given on p. 630. It is instructive to examine how far the properties of the elements may be explained in terms of the arrangements of their electrons.

In accordance with general principles, it is presumed that each atom will, if undisturbed, have the configuration of lowest energy which it can take. Hence if we imagine electrons being successively added, the nuclear charge Ze increasing in step, each fresh electron will go into the state of lowest energy available to it. Pauli's principle only permits one pair of electrons per orbit, as we have seen. In the simplest cases, each shell of electrons is steadily filled in this way and then a new shell is started, to which the shells already completed form a sort of core. As the nuclear charge increases, this core shrinks and its electrons become more tightly bound, but we suppose that each electron preserves the quantum numbers of its state.

The complete system of close on 100 elements involves some complication, but the main features can be discussed with the

[1] W. Pauli, *Zeit. Phys.*, **31**, 765. 1925.

help of the diagram on pp. 628–9. The first two elements, hydrogen and helium, account for the completion of the first or K shell, and this persists unaltered in all subsequent structures, apart from the shrinking and tighter binding mentioned above. At lithium, the L shell is started, but the extra electron is readily lost to form the ion Li^+. Beryllium ($Z=4$) has two easily-detachable or valency electrons, and so on. The L shell is completed with the inert gas neon with its extremely stable electronic configuration. The preceding element is fluorine, which readily accepts an electron to form the negative ion F^- with a stable configuration. Similarly oxygen forms O^{--}. Following neon comes sodium, which is akin to lithium but even more actively electropositive since its odd electron is less strongly held, for the inner shells to a large extent screen the nucleus from this electron.

This period of the table ends with the inert gas argon, although the outer M shell is not yet complete, for it only contains 8 electrons and the Pauli Principle permits $2\times3^2=18$. Nevertheless, the 8 structure in an outer shell appears always to be stable and recurs with all the inert gases except helium. The alkali metals (Li, Na, K, . . .) all follow inert gases and they are increasingly electropositive. In a similar way, the halogens (F, Cl, Br, . . .) precede inert gases, but their electro-negative character becomes less marked along the sequence. This also may be attributed to the more effective screening of the nucleus in the bigger atoms, for on this account an electron, acquired to complete the octet in the outermost shell, is less tightly held, and hence more easily lost again.

The first major complication in the scheme starts after argon, for with potassium one electron goes into the 4th or N shell, although the M shell still lacks its complement of d electrons (with quantum number $l=2$). With calcium, a second $4s$ electron pairs with the first. This s orbit may be pictured as a highly eccentric one, penetrating the shell beneath it and having lower energy than the unfilled $3d$ orbits. At scandium ($Z=21$) building of the M shell is resumed and it is completed at copper ($Z=29$). The elements towards the end of this sequence, and especially the trio Fe, Co, Ni, differ comparatively little from one another and this is to be attributed to the similarity of their electronic structures. Somewhat similar sequences recur in subsequent periods, where the two trios Ru, Rh, Pd and Os, Ir, Pt correspond to Fe, Co, Ni.

An even more remarkable sequence begins at lanthanum ($Z=57$), where the shell wherein most of the additional electrons go is deep within the atom, so that the differences between successive members of the series of elements are very slight. The elements from $Z=57$ to $Z=70$ belong to the Rare Earth Elements, a name which is often taken to include also the kindred elements

scandium ($Z=21$), yttrium (39) and lutecium (71). Recent research indicates that the last elements of the table, starting with actinium (89), form a similar sequence.

The electronic configurations have been worked out mainly on spectroscopic evidence. On the whole the scheme fits the chemical and other physical evidence very well. It may be mentioned that in most cases the removal of one electron from an atom should produce a structure very similar to that of the element preceding it in the table. For example, an ionised alkali metal should have a structure like that of an inert gas. This is in practice true: the spectra of the ionised alkali metals are extremely like those of the inert gases, and also like those of doubly ionised alkaline earths (*e.g.* Ca^{++}).

Magnetic Properties.—It has already been mentioned that measurements on the gyromagnetic effect (p. 296) suggest that ferromagnetism is to be attributed to the action of spinning electrons rather than to orbital motion. According to quantum theory, a pair of spinning electrons interact in such a manner that when they are close to one another, they tend to set antiparallel (*i.e.* with axes parallel but spins oppositely directed), but if the separation increases, this action wanes and they set parallel at somewhat greater distances. Now most of the atomic electrons are paired in their orbits and we can expect actions between neighbouring atoms only when there are unpaired electrons. In the solid state, the most easily detached of the outermost electrons are used in linkages between atoms or, in the case of metals, are conduction electrons. Thus the basic entity in a solid crystal for our purpose is an ion and this is unlikely to have unpaired electrons in its outer layers. In certain cases, however, unpaired electrons may occur in lower layers and this is believed to occur with many of those elements, discussed in the last section, in which an incomplete inner shell exists.

If the spacing of the unpaired spinning electrons is within fairly narrow limits, the spins in adjacent ions line up parallel and the substance develops spontaneous local magnetisation to saturation, the characteristic feature of a ferromagnetic substance (cf. p. 295). This happens with iron, cobalt and nickel in the elementary state and in the correct crystalline form. This ferromagnetism may be lost at high temperature because of thermal agitation or because the element recrystallises to some different form. Compounds and alloys containing one or more of these elements may be ferromagnetic or they may only be paramagnetic.

Manganese is presumably not ferromagnetic in the elementary state because the spins are too close for them to become parallel. In certain alloys, such as the Heusler alloys (p. 286), the spacing is so increased that ferromagnetism appears. Several examples

are now known of ferromagnetic alloys and compounds of two or more non-ferromagnetic elements.

The rare earths as a class are strongly paramagnetic and this is thought to be due to the large separation between the essential spinning electrons. At very low temperatures, where thermal agitation is weak, gadolinium ($Z=64$) is found to be ferromagnetic.

To account for the widespread para- and diamagnetism shown by other substances, contributions must be considered from the orbital electrons, from conduction electrons in solid and liquid metals and from the atomic nuclei themselves. The following books and articles will be found useful for those wishing to pursue the subject further: L. F. Bates, " Modern Magnetism " (Cambridge University Press, 2nd edit., 1948); E. C. Stoner, " Magnetism " (Methuen, 1947); E. C. Stoner, " Ferromagnetism ": articles in *Reports on Progress in Physics* (The Physical Society): Vol. XI (1946–7), p. 43; XIII (1950), p. 83.

Electronic Conduction.—Each atom of a typical metallic element has, in the free state, a small number of electrons which are easily detached, and this accounts, as we have seen, for the characteristic electropositive character of these elements. In the solid state, there are again positive ions in a formation which is too closely packed to allow the outermost electrons to occupy the more eccentric orbits which they might follow in the free state. Hence some of the electrons, of the order of one per atom, are no longer bound to one definite atom. As already discussed (p. 575), Drude sought to explain the characteristic electrical properties of metals in terms of these " conduction electrons " which were represented as forming a kind of electron cloud or gas permeating the metal.

In general, energy must be supplied to release electrons from a block of metal, for if an electron leaves the surface, its repulsion of the other electrons in the vicinity will leave a local positive surface charge and it will therefore have to do work against the resulting attraction if it is to escape. The thermionic and photoelectric effects (see pp. 583 and 603) provide evidence of this energy necessary for escape. There is thus in effect a potential barrier, usually of the order of a few volts, which an electron must surmount if it is to escape.

Thus the conduction electrons may be regarded as attached to the whole piece of metal and not to individual atoms. Just as in the case of electrons constrained by the field of a single atom in Bohr's theory (p. 594), quantum rules are assumed to be valid and there are energy-levels governed by quantum numbers and Pauli's Exclusion Principle (p. 613) applies. Hence each permissible energy-state in a piece of metal may only be occupied by a pair of electrons, oriented with opposed spins. In the undis-

turbed condition at the absolute zero of temperature, all the lowest permissible states will be occupied and there will be a well-defined upper limit.

These close energy-levels fall into *bands* or *zones* and there may be gaps between the bands. A substance which has a fully-occupied band separated from the next band above by a finite gap will behave as an insulator, for the application of a small electric field will be unable to transfer any electron to the upper band of levels and while the lower band is completely full, any electrons moving in one direction are balanced by others moving in the contrary sense. There cannot therefore be any net flow of electricity. On the other hand, when a band is only partly occupied, the field can do work on electrons to transfer them to higher levels and the balance is upset and current flows.

The well-defined gap between bands may be locally blurred or even eliminated by the presence of impurities, which may effectively introduce intermediate levels to which electrons from the lower band can be readily transferred. With rise of temperature, these extra levels become progressively occupied and the substance develops a conductivity which rises with rise of temperature, a characteristic property of semi-conductors.

The emergence of the vigorously growing modern theory of conduction may be roughly dated as starting with the work of Sommerfield,[1] which followed important preliminary work by Fermi, Pauli and others. The reader who wishes to pursue the subject further is referred to: N. F. Mott and H. Jones, "The Theory of the Properties of Metals and Alloys" (Oxford University Press, 1936); A. H. Wilson, "Semi-Conductors and Metals: An Introduction to the Electron Theory of Metals" (Cambridge University Press, 1939); F. Seitz, "The Modern Theory of Solids" (McGraw-Hill, 1940).

The Nucleus.—The small size of the nucleus of the atom and the intense fields which surround it make examination of its structure very difficult. The study of γ-ray spectra shows that there are definite energy-levels in the nuclei emitting them, but this information is limited to a comparatively few nuclei. Certain nuclear properties, including angular momentum (spin) and magnetic moment, are susceptible of indirect measurement, but the greatest amount of information so far has been obtained from the disintegration and transformations of nuclei and indirectly from experimental and observational tests of various theories of nuclear structure.

According to modern views, the nucleus is composed of neutrons and protons, particles of almost equal mass. Taking the isotope ^{16}O of oxygen as having 16 atomic units of mass (16 a.m.u.), the

[1] A. Sommerfield, *Zeit. Phys.*, **47**, p. 1. 1928.

respective masses at rest of the neutron and the proton are 1·00894 and 1·00758. Neglecting for the present the small deviations from unity, it will be seen that an atomic nucleus with N neutrons and P protons will have a mass of $A=N+P$ and a charge, the atomic number, $Z=P$. The neutral atom will also have Z electrons, but an electron has a mass of only 0·00055 a.m.u.

The *nucleons*, as the protons and neutrons of the nucleus are collectively called, must have strong interactions and of the true nature of these forces there are at present only tentative—although sometimes very elaborate—ideas, which are too difficult for the present text. However, it seems clear that the neutrons play an important part in preventing the larger nuclei from exploding under the intense repulsions which are to be expected between protons at such short ranges. The α particle, which is so frequently ejected in radioactive changes, appears to be a particularly stable combination. It has 2 protons and 2 neutrons. Deuterium, or heavy hydrogen (^{2}H), has 1 proton and 1 neutron. In the lighter elements, this ratio is commonly found, those isotopes in which protons account for more than half the mass being unstable. Thus carbon has 6 protons and it has 2 stable isotopes, ^{12}C $(N=6)$ and ^{13}C $(N=7)$, and 3 unstable (radioactive) isotopes, ^{10}C $(N=4)$ with a half-life of only 8·8 sec., ^{11}C $(N=5)$ with 20·5 min., and ^{14}C $(N=8)$ with half-life exceeding 5000 yrs. In the more massive nuclei, the ratio of neutrons to protons is much higher than unity. The following typical examples of stable isotopes show this tendency:—

$^{39}_{19}K(P=19, N=20)$; $^{56}_{26}Fe(P=26, N=30)$; $^{84}_{36}Kr(P=36, N=48)$;
$^{127}_{53}I(P=53, N=74)$; $^{195}_{78}Pt(P=78, N=117)$; $^{208}_{82}Pb(P=82, N=126)$

When a partial disruption of a nucleus occurs, as in radioactive decay, the commonest particles emitted are the α particle, the negative electron and the positron or positive electron. It also sometimes happens that a nucleus absorbs one of the K electrons, leading to subsequent emission of X-radiation when another electron falls into the vacant energy-level. It is not, however, believed that electrons or positrons exist as separate entities in the nucleus: their emission or absorption involves a transformation of a nucleon from one form (neutron or proton) to the other so that electric charge is on the whole conserved.

To induce changes in nuclei, bombardment is carried out by natural particles from radioactive substances as described earlier (*e.g.* p. 540), by artificially accelerated charged particles, by neutrons and by cosmic ray particles. Some of this work will be briefly described.

Artificial Production of Swift Particles.—The earliest nuclear transformation by artificially accelerated particles was achieved

by Cockcroft and Walton,[1] using a steady potential of some 600,000 volts, derived from an alternating source by means of transformer, rectifiers and condensers, to accelerate protons which were generated by a discharge in hydrogen gas. The resulting stream of ions, carrying a current of several microamperes, impinged on selected light elements, such as lithium and boron, and gave a copious supply of α particles which were detected by the scintillation method (p. 516).

Special high-tension generators have been designed to give the high steady potentials necessary for this work. Chief among these is the electrostatic generator of Van de Graaff[2] (Fig. 451). This uses an insulating flexible belt B kept in rapid motion by a motor-driven pulley P. The belt is continuously charged by a discharge (p. 136) from points S connected to a conventional high-tension unit which uses transformers and rectifiers. The belt transfers this charge to the inside of the spherical high-tension terminal T where point-discharge occurs from points C to neutralise it and thus virtually to transfer the charge to the outside of T. The points A, not present in some designs, spray negative charge on to the belt for its downward run, thereby doubling the effective current flowing to the terminal. Potentials exceeding 10 megavolts (10^7 volts) have been obtained by this means, steady to 1 part in several thousand. The accelerating tube for the ions may be housed in the hollow insulating column (not shown in the figure) which surrounds the belt. The reader will appreciate that great care is taken in designing the system to minimise discharge losses and the electrode system is frequently completely shielded and may be in compressed gas to improve the insulation.

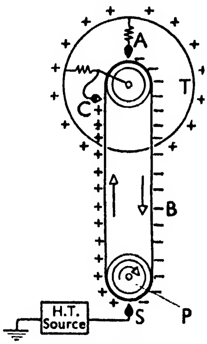


FIG. 451.

Alternative methods have been developed to generate streams of high-velocity particles without very high potentials, using the principle of multiple acceleration. One of the most important of these devices is the *cyclotron* of E. O. Lawrence.[3] In this, the essential operation occurs in the gap between the poles of a large electromagnet. A particle of charge e and mass m, moving with velocity v at right angles to a magnetic field H, experiences a

[1] J. D. Cockcroft and E. T. S. Walton, *Proc. Roy. Soc.*, **137**, p. 229. 1932.

[2] R. J. Van de Graaff, *Phys. Rev.*, **38**, p. 1919 (1931) and **43**, p. 149 (1933). See also the article: R. J. Van de Graaff, J. G. Trump and W. W. Buechner, "Electrostatic generators for the acceleration of charged particles," *Reports on Progress in Physics* (The Physical Society), Vol. XI (1946–7).

[3] E. O. Lawrence and N. E. Edlefsen, *Science*, **72**, p. 376 (1930); E. O. Lawrence and M. S. Livingston, *Phys. Rev.*, **37**, p. 1707 (A) (1931).

constant force transverse to its path and hence has a circular track of radius r given by $mv = eHr$ (p. 471). The angular velocity ω in the path is v/r, so $\omega = eH/m$, and the time required to describe a semicircle is $\pi/\omega = (\pi/H)(m/e)$, depending on H and on the ratio e/m but not separately on the radius r nor on the velocity v.

The vacuum chamber between the pole-pieces of the magnet is nearly filled by two hollow semicircular metal dees (so called from their shape), A and B in Fig. 452. Electrons from a heated tungsten filament below the centre O of the system are repelled, by the negative potential of the filament, into the hollow space within the dees close to O, where they generate by collision positively charged ions of ${}^4_2He$, ${}^1_1H$ or ${}^2_1D$, according to the gas used—helium, hydrogen or deuterium.

The potential difference between A and B is alternating. Assuming that B is momentarily at the higher potential, a positive ion from near C will be driven into A with appreciable velocity. Within A, the electric field is zero, so that the ion moves, under the action of the magnetic field, in a circle and arrives at D after a time $(\pi/H)(m/e)$. The periodic time of the alternating potential is chosen to be twice this, so that when the ion arrives at D the electric field between the dees is reversed, and once again the ion is accelerated, entering B with a greater velocity than it entered A. The process is repeated at E, F, etc., the velocity and increase of radius of path at each passage across the gap persisting until the outer edge of the dee is reached, when the velocity is Hae/m, a being the radius of the dee. For example, with a radius of 40 cm. and a field of $1{\cdot}3 \times 10^4$ gauss, a deuteron acquires a velocity corresponding to a fall through a potential difference of $6{\cdot}4 \times 10^6$ volts.

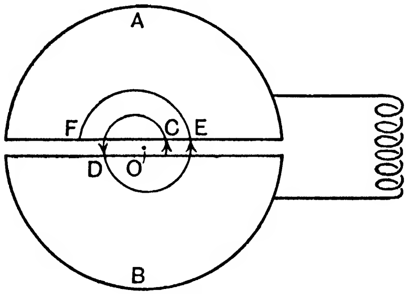


FIG. 452.

When the ion reaches the boundary of the dee, at a point such as A or B (midway between the gaps) it may be withdrawn by attraction towards a plate at negative potential, the dees being at this moment at zero potential. Ions which make their first entry into a dee when the electric field is not at its maximum will receive less than the maximum increase in velocity at each traverse of the gaps and will therefore have to make more circuits before reaching the edge. They thus stand more chance of making a collision in the rarefied gas of the chamber, but if they survive this, they ultimately acquire the same velocity on reaching the

periphery. The total path of each ion may be many hundreds of feet.

Among the other accelerators which use magnetic fields, mention may be made of the *synchro-cyclotron*, in which the frequency of alternation is periodically varied to compensate for the increase in mass of the particles at high speeds (cf. p. 546); and the *betatron*, first successfully operated in 1941 by Kerst, in which electrons are accelerated by a magnetic field which increases with the time. In this case the electrons are accelerated by the E.M.F. induced by the changing magnetic flux. The betatron may be used to generate very penetrating X-rays by placing a suitable target (*e.g.* a platinum wire) in the path of the high-speed electrons.

There are also *linear accelerators*, in the simplest form of which a number of cylindrical metal tubes in line are so coupled to a high-frequency oscillator that an ion may be accelerated every time it passes from one section to the next, the reverse changes in potential occurring while the ion is in the field-free space inside a section.

Some of these accelerators generate particles with energies as high as 500 or even 1000 Mev. For further details, the following articles may be consulted: O. R. Frisch, " Artificial Acceleration of Atomic Particles," *Nature*, **168**, p. 849 (17 Nov. 1951); D. W. Fry and W. Walkinshaw, " Linear Accelerators," *Reports on Progress in Physics*, Vol. XII (1948–9), p. 102; J. H. Fremlin and J. S. Gooden, " Cyclic Accelerators," *ibid.*, Vol. XIII (1950), p. 295.

If deuterons from a cyclotron or other accelerator impinge on a suitable target of a light element, a powerful beam of neutrons is obtained, apparently derived from the break-up of the deuterons. The speed of the neutrons depends on the target material and the speed of the original deuterons.

Photographic Emulsions for Nuclear Study.—It was shown by W. Michl[1] that α particles can affect a photographic plate, and Marietta Blau extended the technique first to fast protons[2] and then to cosmic rays.[3] After development, the tracks of swift charged particles through the emulsion are visible under the microscope. The production of the primary specks of silver which form the " latent image " of photography may be regarded essentially as an ionising effect and the process of development leads to a deposit of silver round these primary specks. Thus there is a fairly close analogy between these photographic tracks and those produced by condensation on ions in a cloud chamber. The particles lose their energy more rapidly in the solid emulsion

[1] W. Michl, *Wien. Ber.*, **121**, p. 1431 (1912) and **123**, p. 1955 (1914).
[2] Marietta Blau, *Zeit. Phys.*, **34**, p. 285. 1925.
[3] M. Blau and H. Wambacher, *Akad. Wiss Wien.*, **146**, 2a, pp. 469 and 623. 1937.

than in a gas, so that the scale of the tracks is much smaller and they must be examined under high-power microscopes or be much enlarged. In the hands of Powell and his collaborators at Bristol, and others, the photographic method has proved of great value. For example, balloons can be sent to some 30 km., carrying a stack of specially coated photographic plates, suitably wrapped. After a pre-arranged time, the load is automatically cut adrift from the balloon and descends by parachute. The stack of plates is usually recovered intact and the plates are then processed and examined under microscopes.

The photographic observations have shown that the primary radiation reaching the earth's atmosphere from space is mostly composed of fairly massive atoms or ions moving at very great speed. These disintegrate, sometimes into very many fragments, on collision with atoms in the upper terrestrial atmosphere and these fragments, together with radiation and other fragments liberated in turn from other atoms lower in the atmosphere, constitute the cosmic radiation experienced at the earth's surface.

Mesons.—In 1935 Yukawa,[1] in developing a theoretical explanation of the forces holding together the nucleons in a nucleus, postulated a particle, the meson, with a mass some 200 times that of an electron. Some time later, cloud-chamber studies of cosmic rays revealed tracks which corresponded, from density of ionisation and curvature in a magnetic field, to particles of about this mass, and intensive study using the photographic emulsion technique has disclosed that mesons (formerly also called mesotrons) are commonly emitted when a nucleus is disrupted and that they may have charges of electronic magnitude but of either sign or they may be uncharged. These primary mesons are written as π^+, π^- or π^0 according to their charge.

Charged π^--mesons have a mass of about 280 m_e, where m_e denotes the mass of the electron, and they decay spontaneously in a life-time of about $2{\cdot}6\times10^{-8}$ sec.,[2] giving rise to a lighter meson, the μ-meson, of similar charge but mass about 215 m_e. The π^+-meson is repelled by atomic nuclei, but when passing through matter, the π^--meson has a high chance of entering an atom before it decays. When the spontaneous decay $\pi \rightarrow \mu$ occurs, some undetected neutral body of high velocity and small rest-mass is postulated, in order to balance momentum and energy. Such a *neutrino* must have a mass less than $0{\cdot}1\ m_e$.

Charged μ-mesons constitute the bulk of the " hard " or penetrating component of cosmic rays. The μ-meson in turn decays, after a mean life of some 2×10^{-6} sec., to form a similarly-charged electron, probably also giving rise to two neutrinos. Neutral

[1] H. Yukawa, *Proc. Phys. Math. Soc.*, Japan, **17**, p. 48. 1935. (In English.)
[2] C. E. Wiegand, *Phys. Rev.*, **83**, p. 1085. 1951.

π-mesons have masses of the order 270 m_e and extremely short life, of order 10^{-13} sec., and disrupt to give γ-radiation which in turn gives rise to cascade showers of electron-pairs (positive and negative) and photons, these accounting for the " soft " component of cosmic radiation.

The primary cosmic radiation contains nuclei of elements up to about iron ($Z=26$). These nuclei, on making collisions with molecules in the very high atmosphere, at heights mostly over 60,000 ft., disrupt more or less completely, each giving rise to a shower of nucleons and π-mesons and the nucleons in turn give rise to more π-mesons by collision farther down in the atmosphere. The μ-mesons observed lower still arise by decay of the π-mesons.

More massive mesons, known as τ-mesons, with mass of the order 1000 m_e, have been detected on a number of occasions and still more types may exist. The terms ρ-meson and σ-meson were provisional names referring to the originators of certain types of track: the σ-mesons produced observable disintegrations and the ρ-mesons did not. In fact σ-mesons are now believed to be μ^--mesons, and ρ-mesons to be a mixture of μ^+, μ^- and perhaps some π^+-mesons.

For a fuller account the reader may refer to the article " Mesons " by C. F. Powell in *Reports on Progress in Physics* (The Physical Society), Vol. XIII (1950), p. 350.

Nuclear Fission.—It has been mentioned (p. 511) that mass and energy are regarded nowadays as equivalent and the stability of nuclei may be discussed from this point of view. Thus the nucleus of the commonest isotope of uranium, $^{238}_{92}U$, has 92 protons with a combined mass of $92\times1{\cdot}00758=92{\cdot}6973$ a.m.u. (atomic mass units) plus 146 neutrons, of $146\times1{\cdot}00894=147{\cdot}3053$ a.m.u. Allowing $92\times0{\cdot}000549=0{\cdot}0505$ a.m.u. for the 92 extra-nuclear electrons, the total mass of the constituents of the neutral atom is 240·053 a.m.u., while the observed mass of the atom is 238·136. Thus if such an atom could be formed from its constituent particles, energy equivalent to 1·917 a.m.u. would be radiated, some $1{\cdot}917\times1{\cdot}66\times10^{-24}\times9\times10^{20}=2{\cdot}9\times10^{-3}$ erg, since 1 a.m.u.$=1{\cdot}66\times10^{-24}$ erg. The other factor is c^2 (p. 511). This energy is equivalent to some 1800 Mev.

Despite this mass-deficit or binding energy, which precludes spontaneous disintegration into its constituents, the uranium nucleus might conceivably break into fragments of even greater total binding energy. Thus in the various steps in a radioactive series, the total mass-defect (including where appropriate that of emitted particles) continually increases. Now the binding energy per nucleon is much higher for elements of moderate atomic number, where the neutron : proton ratio is roughly unity, than it is for the heavy atoms. In consequence, the fission of a heavy

nucleus into two or more parts of comparable mass would in many cases be accompanied by the liberation of energy.

We know now that Hahn, Meitner and Strassmann obtained fission of uranium in 1934 by neutron bombardment of uranium, but at the time the observed rise in radioactivity was wrongly attributed to the creation of elements of higher atomic number. The incorrect hypothesis was disproved by Curie and Savitch in 1939, and in the same year Hahn and Strassmann demonstrated the true nature of the process. The uranium nucleus first absorbs the neutron, forming the unstable isotope ${}^{239}_{92}U$, and this then disintegrates into nuclei of unequal but comparable mass, not always in the same way. Many elements have been detected in the fission products. They all have neutron : proton ratios nearer to unity than for the original nucleus and hence several fresh neutrons may be liberated by each fission.

${}^{238}_{92}U$ only gives fission when struck by a fast neutron, although it will absorb neutrons of lower velocity, without fission. The istope ${}^{235}_{92}U$, which constitutes 0·7 per cent. of natural uranium, will give fission with both fast and " slow " neutrons. Neutrons may be slowed down by passing them through material containing light nuclei, hydrogen, deuterium (heavy hydrogen, ${}^{2}_{1}H$) and carbon being especially suitable. Fast neutrons can also give rise to fission in thorium.

The liberation of fresh neutrons by a fission makes possible a *chain reaction*, which can be controlled in an *atomic* (or more appropriately, *nuclear*) *pile* or *nuclear reactor*, containing uranium rods fitted into a matrix of graphite, and adjustable rods or screens of cadmium or other material capable of absorbing neutrons. The graphite serves as moderator, that is to say, it slows the neutrons to the low velocities which are especially effective in producing the fission of ${}^{235}U$. Some neutrons generated by fission escape or are absorbed; for a steadily maintained reaction, on an average one neutron from each fission must induce a further fission and the rods or screens may be adjusted to attain this condition. The large thermal energy which may be liberated can be removed by a cooling agent, which may be water or a suitable gas flowing in tubes in and around the pile.

Within the reactor there is an intense flux of neutrons and many nuclear reactions can be produced by placing samples inside the pile. The main uranium isotope ${}^{238}U$ itself provides an important example, giving first a short-lived radioactive isotope:

$$ {}^{238}_{92}U + {}^{1}_{0}n \rightarrow {}^{239}_{92}U + \gamma \text{ rays.} $$

This isotope decays by β-emission, with a half-life of 23 min., to give a new element of atomic number 93, *neptunium*:

$$ {}^{239}_{92}U \rightarrow {}^{239}_{93}Np + {}^{0}_{-1}\beta \text{ (electron).} $$

In this process, a neutron is presumed to change into a proton. Neptunium has a half-period of 23 days, and gives rise to *plutonium*:

$$^{239}_{93}\text{Np} \rightarrow {}^{239}_{94}\text{Pu} + {}_{-1}^{0}\beta + \gamma \text{ rays.}$$

Plutonium has a long half-value period, some 24,000 yrs. It decays by α-emission and gives rise to $^{235}_{92}$U, the rarer uranium isotope. It is therefore shown on the chart of p. 539 as a progenitor of the actinium series, although in the course of geological time it, and its parent element, have decayed completely. The very small amount of plutonium found in nature has presumably been generated from uranium by neutron action and does not represent any residuum from the initial elements which gave rise to this family.

Plutonium is readily fissionable by neutrons and when it splits it liberates further neutrons, so that it may be used for a chain reactor. If a block of plutonium has sufficient volume and compactness to ensure the utilisation of an adequate proportion of the emitted neutrons, since there are always sufficient stray neutrons to initiate the process, the mass will spontaneously explode. Thus one way to produce and fire a nuclear bomb is to bring into effective contact two or more portions of plutonium, each separately below the critical size for spontaneous disintegration. ^{235}U may also be used.

Artificial Isotopes.—By means of artificial sources of swift particles, and particularly of neutrons, a very great number of nuclear reactions have been produced, and many of these result in the production of new isotopes. Some 300 different isotopes have been recognised in nature, but already far more have been produced artificially than previously existed.

The majority of the new isotopes are elements already known, but some quite new elements have been produced. These include some elements previously missing from the Periodic Table (*e.g.* astatine, a halogen, Z=85, and francium, an alkali, Z=87) and several trans-uranic elements. The latter are:—

93 Neptunium (Np); 94 Plutonium (Pu); 95 Americium (Am);
96 Curium (Cm); 97 Berkelium (Bk); 98 Californium (Cf).

All these are radioactive. They are chemically closely allied and are believed to form a series somewhat like the rare earths (cf. pp. 614–5).

Artificial isotopes have already found many uses, especially the strongly radioactive ones. Detection of radiations from these is such a delicate method of indicating their presence that a very small quantity of a radioactive isotope may be used as a *tracer*, enabling the progress of the sample containing it to be followed. The rest of the sample, if composed of another isotope, will

behave chemically and physically in exactly the same way, or almost so, and the two isotopes therefore pursue the same course. In this way radioactive forms of carbon, sodium, phosphorus, iodine and many other elements are used to indicate the utilisation of foodstuffs by plants and animals and the exchanges of elements between tissues. Among many engineering applications, the transfer of traces of material in sliding friction may be followed. In some cases the flow and mixing of chemicals may be followed by detectors which are outside the system studied. Tracer techniques are also used in fundamental studies of chemical reactions.

NOTES ON ATOMIC TABLES AND CHARTS

ATOMIC CHARTS, pp. 628–9

For each known element, the following information is given:

1. Atomic number (Z) (see p. 496).
2. Name and chemical symbol.
3. The approximate masses of all its well-established isotopes, distinguishing between stable and unstable (i.e. radioactive) forms (see pp. 503, 539).
4. The arrangement of the electrons in the neutral atom (p. 610).

Additional information about percentage abundance of naturally occurring isotopes and the half-value periods (p. 526) of the radioactive isotopes will be found in *Tables of Physical and Chemical Constants* by G. W. C. Kaye and T. H. Laby (Longmans, Green & Co., Ltd.).

The electronic chart not only gives the distribution of the Z electrons of the neutral atom of the element of atomic number Z, but also indicates qualitatively the shrinking in size of the various levels as the nuclear charge (Ze) increases. Completed shells are indicated by shading and completed sub-shells by dots. Some of the assignments to levels are provisional only.

THE PERIODIC TABLE, p. 630

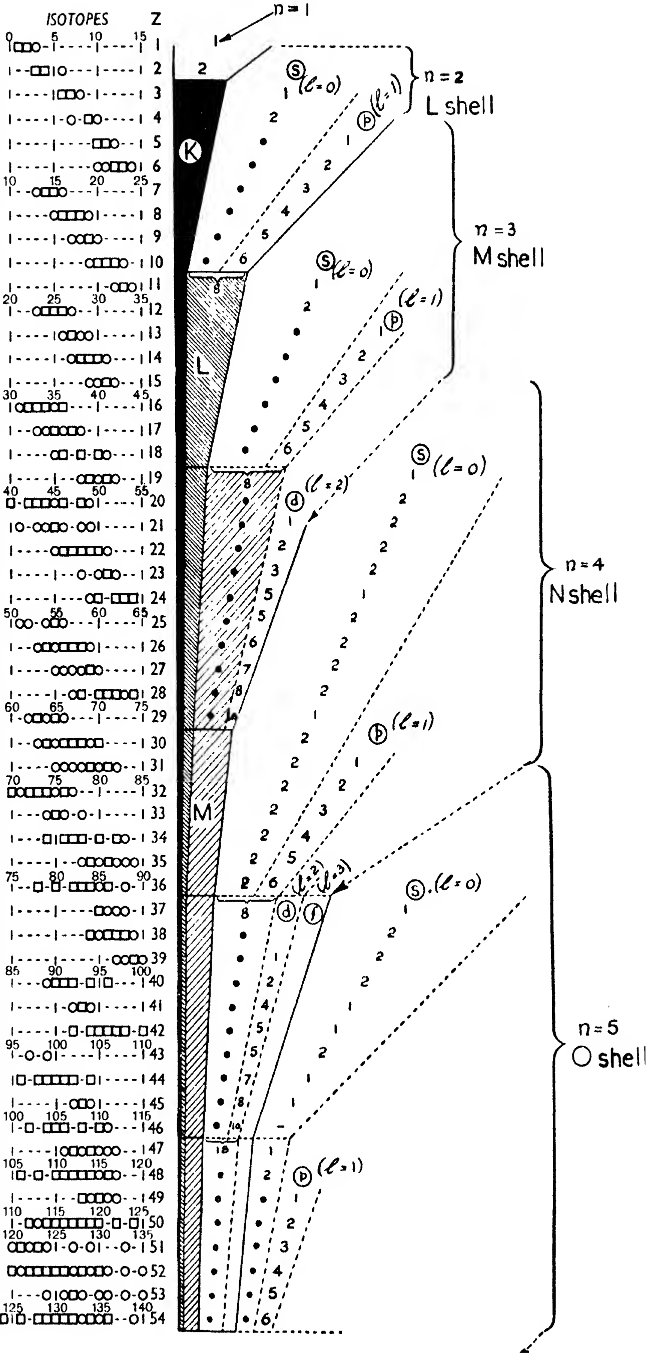
The arrangement of the elements in this table follows broadly the scheme due to Mendeléef. Each entry gives the atomic number, chemical symbol and atomic weight of the element. Where the element is only known in the form of artificial isotopes, the approximate "atomic weight" (rounded off to the nearest integer) of the most stable isotope (the one of greatest half-value period) is given in brackets, thus:

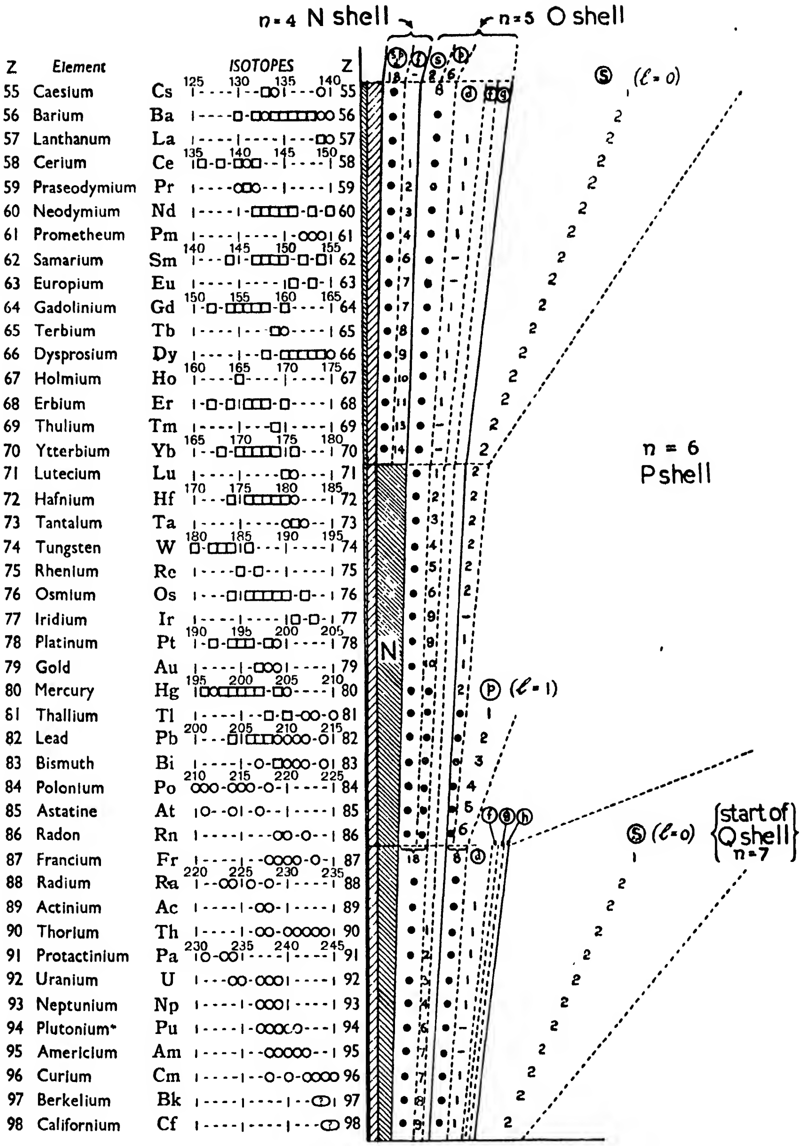
43 Tc
[99]

For a discussion of the relationship of the periodicity in properties of the elements to the electronic structure displayed on pp. 628–9, see p. 613. Each structure may be regarded as derived from the one preceding by the addition of an electron. Where this extra electron does not enter the outermost shell, the symbols *, † are used to indicate whether the shell enriched is one place or more below this outer one. Some of these cases, especially at the end of the table, are conjectural.

Until recently, the elements thorium (*90* Th), protoactinium (*91* Pa) and uranium (*92* U) were generally assigned respectively to groups IV, V, VI. The Table shows them in brackets in this position; the more modern practice is to place them, as shown, in an actinide series analogous to the lanthanide rare-earth series.

Z	Element	
1	Hydrogen	H
2	Helium	He
3	Lithium	Li
4	Beryllium	Be
5	Boron	B
6	Carbon	C
7	Nitrogen	N
8	Oxygen	O
9	Fluorine	F
10	Neon	Ne
11	Sodium	Na
12	Magnesium	Mg
13	Aluminium	Al
14	Silicon	Si
15	Phosphorus	P
16	Sulphur	S
17	Chlorine	Cl
18	Argon	A
19	Potassium	K
20	Calcium	Ca
21	Scandium	Sc
22	Titanium	Ti
23	Vanadium	V
24	Chromium	Cr
25	Manganese	Mn
26	Iron	Fe
27	Cobalt	Co
28	Nickel	Ni
29	Copper	Cu
30	Zinc	Zn
31	Gallium	Ga
32	Germanium	Ge
33	Arsenic	As
34	Selenium	Se
35	Bromine	Br
36	Krypton	Kr
37	Rubidium	Rb
38	Strontium	Sr
39	Yttrium	Y
40	Zirconium	Zr
41	Niobium	Nb
42	Molybdenum	Mo
43	Technetium	Tc
44	Ruthenium	Ru
45	Rhodium	Rh
46	Palladium	Pd
47	Silver	Ag
48	Cadmium	Cd
49	Indium	In
50	Tin	Sn
51	Antimony	Sb
52	Tellurium	Te
53	Iodine	I
54	Xenon	Xe





(For general explanation, see p. 627)

Stable isotopes □

Radioactive isotopes ○

Where both stable and radioactive isotopes of an element exist having the same mass, the stable form is indicated.

PERIODIC TABLE

GROUP:	I		II		III		IV		V		VI		VII		VIII			0
PERIOD	a Alkali Metals	b	a Alkaline Earth Metals	b	a	b	a	b	a	b	a	b	a	b Halogens	Transition Elements			Inert Gases
1		1 H 1·0080																2 He 4·003
2	3 Li 6·940		4 Be 9·013			5 B 10·82		6 C 12·010		7 N 14·008		8 O 16·000		9 F 19·00				10 Ne 20·183
3	11 Na 22·997		12 Mg 24·32			13 Al 26·98		14 Si 28·09		15 P 30·975		16 S 32·066		17 Cl 35·457				18 A 39·944
4	19 K 39·100	*29 Cu 63·54	20 Ca 40·08	30 Zn 65·38	*21 Sc 44·96	31 Ga 69·72	*22 Ti 47·90	32 Ge 72·60	*23 V 50·95	33 As 74·91	*24 Cr 52·01	34 Se 78·96	25 Mn 54·93	35 Br 79·916	*26 Fe 55·85	*27 Co 58·94	*28 Ni 58·69	36 Kr 83·80
5	37 Rb 85·48	47 Ag 107·880	38 Sr 87·63	48 Cd 112·41	*39 Y 88·92	49 In 114·76	*40 Zr 91·22	50 Sn 118·70	*41 Nb 92·91	51 Sb 121·76	*42 Mo 95·95	52 Te 127·61	43 Tc [99]	53 I 126·91	*44 Ru 101·7	*45 Rh 102·91	*46 Pd 106·7	54 Xe 131·3
6	55 Cs 132·91	*79 Au 197·2	56 Ba 137·36	80 Hg 200·61	*57 La †58 Ce †59 Pr ... ... *70 Yb *71 Lu	Lanthanide Series of Rare Earths 81 Tl 204·39	*72 Hf 178·6	82 Pb 207·21	*73 Ta 180·88	83 Bi 209·00	*74 W 183·92	84 Po 210	*75 Re 186·31	85 At [210]	*76 Os 190·2	*77 Ir 193·1	78 Pt 195·23	86 Rn 222
7	87 Fr [223]		88 Ra 226·05		*89 Ac †90 Th †91 Pa †92 U ... ... †98 Cf	Actinide Series	(†90 Th 232·12)		(†91 Pa 231)		(†92 U 238·07)							

LANTHANIDE SERIES	*57 La 138·92	†58 Ce 140·13	†59 Pr 140·92	†60 Nd 144·27	†61 Pm [145]	†62 Sm 150·43	†63 Eu 152·0	*64 Gd 156·9	ACTINIDE SERIES	*89 Ac 227	†90 Th 232·12	†91 Pa 231	†92 U 238·07	†93 Np [237]	†94 Pu [242]	†95 Am [243]	*96 Cm [243]
		†65 Tb 159·2	†66 Dy 162·46	†67 Ho 164·94	†68 Er 167·2	†69 Tm 169·4	†70 Yb 173·04	*71 Lu 174·99			†97 Bk [245]	†98 Cf [246]					

For notes on the Table, see p. 627. * denotes entry of electron to an inner shell, † to a still deeper shell. [] enclose (approximate) masses of most stable (longest lived) known isotopes of artificially produced radioactive elements.

EXAMPLES

Examples I

1. Two short magnets, with their axes horizontal, and perpendicular to the magnetic meridian, are placed with their centres 30 centimetres east and 20 centimetres west respectively of a compass-needle. Compare the moments of the magnets if the needle remains undeflected, and show how to derive the formula employed in the calculation. [$M_E/M_W=27/8$.]

2. Find an expression for the potential at a point in the field due to a very short magnet of known moment. Hence deduce the components of the field intensity at the point, along and perpendicular to the line joining it to the centre of the magnet. (L.U., B.Sc. Internal.)

3. Find expressions for the turning effect and for the attraction of one small magnet on another. (L.U., B.Sc. Hon. Internal.)

4. Show how mechanical principles are not violated by the fact that the couple exerted by one magnet on another is, in general, not the same as that of the second on the first. (L.U., B.Sc. Hon. Internal.)

5. Two magnets of the same length l are placed with their axes parallel and their centres at a considerable distance R from a point P, one with its axis passing through P, the other with P on the line through its centre perpendicular to its axis. Find how the magnets must be oriented, and what must be the relation between their moments in order that the magnetic field at P due to them may be independent of powers of $\frac{l}{R}$ lower than the fourth. (L.U., B.Sc. Ext.) [Magnets oriented as N_1S_1, S_2N_2; $M_1/M_2=3/8$.]

6. Find the magnitude and direction of the magnetic field due to a small magnet of moment 30, at a point situated on a line passing through the middle of the magnet and at an angle of 60 degrees with its axis, the point being at a distance of 5 cm. from the magnet. [$3\sqrt{7}/25$ at 100° 54′ to axis of magnet.]

7. Describe, giving all necessary correcting terms, how you would determine, in absolute measure, the horizontal component of the earth's magnetic field. (L.U., B.Sc. Internal.)

8. Describe some method of comparing H and V, the horizontal and vertical components of the earth's magnetic field.

Show that the ratio $H \div V$ would be equal to $\frac{\cot\theta}{2}$, where θ is the latitude, if the magnetic field were due to a magnet at the centre of the earth with its axis pointing north and south. (L.U., B.Sc. Ext.)

9. Describe the method adopted to determine the variation of the horizontal component of the earth's magnetic field, and find an expression for the deflection produced in terms of the variation of the component. (L.U., B.Sc. Ext.)

10. Give an outline of the determination of H—the horizontal magnetic field of the earth—by employment of a bar magnet and a magnetic needle.

How is the effect of the length of the bar magnet eliminated by observing deflections of the needle by the bar magnet at two different distances of the latter ? (L.U., B.Sc. Internal.)

11. Describe how the variation of the intensity of the earth's magnetic field may be continuously recorded.

12. A magnet weighing 15 grammes has a small straight stem (length=4 mm.) fixed centrally at right angles to the magnetic axis. The whole is suspended by a silk fibre attached to the upper end of the stem, at a place where the dip is 60°. Calculate the angle which the magnet makes with the horizontal, its magnetic moment being 100 units. (L.U., B.Sc. Internal.) [$\tan\theta = 100H\sqrt{3}/(6g+100\ H)$.]

Examples II

1. What is the reason for using as small a suspended magnet as possible in the tangent galvanometer?

Describe an arrangement of coils by which the necessity for a very small magnet is removed, stating the reasons.

2. Two similar coils of wire, having a radius of 7 cm. and 60 turns, have a common axis and are 18 cm. apart. Find the strength of magnetic field (*a*) at the centre of either coil, and (*b*) at a point on their common axis midway between them. [(*a*) 5·642 per amp.; (*b*) 2·493 per amp.]

3. Find the resistance of a cubic centimetre of copper (*a*) when drawn out into a wire of diameter 0·32 mm. and (*b*) when hammered into a flat sheet of thickness 1·2 mm. the current flowing perpendicularly through the sheet from one face to the other. (Specific resistance $= 1{\cdot}59 \times 10^{-6}$.) [(*a*) 2·549; (*b*) $2{\cdot}29 \times 10^{-8}$ ohm.]

4. State the laws governing the distribution of current in a network of wires.

A battery of 6 volts E.M.F. and 0·5 ohm internal resistance is joined in parallel with another of 10 volts E.M.F. and 1 ohm internal resistance, and the combination used to send current through an external resistance of 12 ohms. Calculate the current through each battery. (L.U., B.Sc. Ext.) [2·865; $-2{\cdot}270$ amp.]

5. State Kirchhoff's Laws of distribution of electric currents in networks of conductors, and justify by means of them or otherwise the method of determining the resistance of a voltaic cell by placing it in one arm of a resistance bridge. (L.U., B.Sc. Ext.)

6. Define the ampere, and find the direction and intensity of the force on a circular coil of n turns wound close together through which a current of A amperes is flowing due to a magnet whose poles lie on the axis of the coil. (L.U., B.Sc. Ext.)

7. The reactions within a cell generate electrical energy at the rate of 1 watt per ampere; a current of 10 amperes is being generated with the result that energy is dissipated within the cell in the form of heat at the rate of 1 watt. What is the difference of potential between the terminals of the cell, also what is the internal resistance of the cell? (L.U., B.Sc. Internal.) [0·9; 0·01.]

8. Show how to calculate the current through the galvanometer in the Wheatstone bridge arrangement of conductors when nearly balanced. (L.U., B.Sc. Hon. Internal.)

Examples III

1. Describe how you would determine accurately a very small resistance.

2. Describe the moving coil galvanometer, giving its advantages and disadvantages compared with the suspended needle galvanometer.

3. Describe carefully how you would use a potentiometer for measuring currents. How would you adapt it for use with large and small currents respectively? (L.U., B.Sc. Internal.)

4. Describe the construction of the moving coil galvanometer, and explain how, with the addition of a shunt, it can be used as an ammeter for large currents. (L.U., B.Sc. Ext.)

5. What difficulties are met with in measuring a very small resistance by the Wheatstone's bridge method? Describe a method of comparing low resistances. (L.U., B.Sc. Internal.)

6. Describe some form of ammeter in which the reading is proportional to the square of the current.

Examples IV

1. Why does a sharp point attached to a conductor prevent a high potential being obtained, while a knob has no such effect.

Describe some practical application of the above action of points.

2. Give an account of the method employed by Cavendish and Maxwell to prove that the law of inverse square of the distance holds in electrostatics. How can the degree of accuracy attained in such experiments be estimated? (L.U., B.Sc. Internal.)

3. Define the term *potential*, as applied to conductors in electrostatics.

Show that the potential must be the same at all points in the air space completely surrounded by a conductor. (L.U., B.Sc. Ext.)

4. What is an electrical image?

A point charge is placed 3 cm. in front of an infinite plane conductor. Show that the total induced charge on the portion of the plane which is contained by the circumference of a circle of radius 4 cm., and whose centre is the foot of the perpendicular let fall from the point charge on to the plane is numerically $\frac{2}{5}$ of the point charge. (L.U., B.Sc. Hon. Internal.)

5. Show that the amount of energy per unit volume of an electric field at any point P in the field is $\frac{kR^2}{8\pi}$, where k is the specific inductive capacity and R the electric intensity at P.

6. Show that the electric force close to a charged surface is normal to the surface and equal to $4\pi\sigma$, where σ is the surface density of charge.

Show also that the force acting on unit area of the surface is $2\pi\sigma^2$ along the direction of the normal. (L.U., B.Sc. Ext.)

7. Describe the application of the method of images to the solution of electrostatic problems.

8. Show that in passing from one dielectric to another, electric lines of force undergo a change in direction.

9. Find the force with which a point charge of electricity and a conducting sphere attract each other when the sphere is earthed. $[q^2rd(d^2-r^2)^2.]$

10. What is meant by an electrical image? A charge of electricity $+q$ is situated at a distance l from a large earthed plane conducting sheet. Find the distribution of the induced charge in terms of the distance from the point. (L.U., B.Sc. Ext.)

11. Find an expression for the force per square centimetre of surface on a conductor due to its charge. What charge must there be upon a soap bubble of radius $1\frac{1}{2}$ cm. if the air pressure is the same inside and outside of the bubble, assuming the surface tension to be 27? (L.U., B.Sc. Hon. Internal.) [95·7 e.s.u.]

12. Show that if there is no force inside a uniformly charged spherical surface the law of force between two charges e_1 and e_2 is $\frac{e_1e_2}{r^2}$, where r is the distance between the two points at which e_1 and e_2 are situated. (L.U., B.Sc. Hon. Ext.)

13. Show how the induced electrification distributes itself on a conducting sphere placed in a uniform field. (L.U., B.Sc. Hon. Internal.)

14. Explain the method of electrical images for the solution of problems in electrostatics.

Find the distribution of electricity produced on a conducting sphere insulated without charge, when a point charge is placed near it. (L.U., B.Sc. Hon. Ext.) $[\sigma=qr\left(1-\frac{d^2}{r^2}\right)/4\pi PM^3+q/4\pi rd.]$

Examples V

1. Two spheres of radii 5 and 10 cm., respectively have equal charges of 50 units each. They are then joined by a thin wire so that their charges are shared between them. Calculate the total energy before and after sharing. What becomes of the difference of energy? (L.U., B.Sc. Internal.) [375 ergs; $333\frac{1}{3}$ ergs.]

2. Describe some form of absolute electrometer and give the theory of its action. (L.U., B.Sc. Hon. Internal.)

3. Define accurately the term *capacity* of a condenser, and show which has the greater capacity, the inside or the outside coating of a Leyden jar.

Find the capacity of a sphere of 15 cm. diameter inside which there is an earthed concentric sphere of 10 cm. diameter. (L.U., B.Sc. Ext.) [22·5 e.s.u.]

4. The ends of a metal tube of radius r_1 project into two larger tubes of radii r_2 and r_3, the radii being small compared with the lengths of the tubes, and the axes being all in the same line. Find the force in dynes on the inner tube when the potentials of the three conducting surfaces are v_1, v_2, and v_3 respectively. (L.U., B.Sc. Hon. Ext.) $[v_1-v_2)^2/9{\cdot}10 \log_{10} (r_2/r_1)-(v_1-v_2)^2/0{\cdot}21 \log_{10}/r_3/r_1.]$

5. A wire 1 mm. in diameter is stretched along the axis of a conducting cylinder whose internal radius is 1 cm. Calculate the capacity of the structure per unit length. (L.U., B.Sc. Ext.) [0·1669 e.s.u. per cm.]

6. How would you determine the specific inductive capacity of a solid substance, being given a slab of the material in question?

7. Given a standard condenser of 1/3 microfarad, how would you use it to determine a very small capacity (*e.g.*, about the ten-thousandth part of the standard)?

8. Find the capacity, per square centimetre, of a condenser formed of two parallel conducting planes of infinite extent when the intervening gap consists of 8 mm. of air and 4 mm. of a substance whose dielectric constant is 5. (L.U., B.Sc. Ext.) [0·0904 e.s.u. per sq. cm.]

9. Deduce an expression for the electrostatic capacity of two coaxial metallic cylinders separated by a layer of air. Investigate the effect of inserting between the cylinders a coaxial cylindrical shell of a dielectric substance of thickness less than that of the layer of air. (L.U., B.Sc. Hon. Ext.)

10. Describe some form of quadrant electrometer and deduce a formula for use with it. (L.U., B.Sc. Ext.)

11. Two parallel conducting plates are maintained with a constant difference of potential between them. Find the ratio of the attractions between the plates when air is the only medium separating them, and when a sheet of non-conducting material whose thickness is two-thirds of the distance between the plates and whose dielectric constant is 6, is inserted. [16/81.]

12. A cable consisting of a solid conductor of 6 mm. diameter is surrounded by two layers of insulating material, separated by a thin conducting layer, the inner having a thickness of 3 mm. and a dielectric constant 7, and the other a thickness of 4 mm. and dielectric constant 5. Outside this is an earthed conducting sheath. Find the ratio of the falls of potential in the two insulating layers. On gradually raising the potential difference between the inner conductor and the earthed sheathing, which of the layers will first break down in insulation, assuming the electric strength of the two insulating materials to be the same? (L.U., B.Sc. Hon. Internal.) [0·969; inner.]

13. Define the terms *electrical potential*; *capacity of a condenser*.

A condenser is made up of two concentric spheres of thin metal, radii 5 cm. and 8 cm. respectively, and there are no other conductors in the neighbourhood. If a charge +5 e.s.u. is given to the inner sphere, and +15 e.s.u. to the outer sphere, calculate the potentials of the spheres and state how the charge is distributed on the outer sphere. (L.U., Scholarships.) [2·875 inner; 2·5 outer; −5 e.s.u., +20 e.s.u.]

14. Prove Gauss's theorem in electrostatics.

Find the capacity of unit length of a cylindrical condenser of which the conductors have radii 2·5 cm. and 4·5 cm. respectively, and the dielectric consists of two layers whose cylinder of contact is 3·5 cm. in radius; the inner layer having a dielectric constant 4 and the outer layer a dielectric constant of 6. (L.U., B.Sc. Internal.) [3·98 e.s.u. per cm.]

15. Obtain an expression for the stress on the surface of a charged conductor in terms of the electric intensity at the surface and the dielectric constant of the material outside it.

A condenser consists of two coaxial cylinders of radii 5 and 10 cm., separated by a medium of dielectric constant 3. Calculate the pull on each of the charged surfaces per unit area when the potential difference between the cylinders is 200 volts. (L.U., B.Sc. Internal.) [0·00442 dyne cm.$^{-2}$ on inner; 0·00110 dyne cm.$^{-2}$ on outer.]

Examples VI

1. Describe some method of determining the specific resistance of an electrolyte.

2. State the evidence, derived from the facts of electrolysis, that electricity is atomic in structure. (L.U., B.Sc. Internal.)

3. Describe the phenomenon of electrolytic conduction and explain how ionic velocities have been found.

4. Give a short account of the ionic theory of electric conduction in electrolytes, and explain why a difference in potential should be expected when diffusion of a salt takes place.

5. Discuss the relation between the E.M.F. of a Daniell cell and the chemical changes which take place in it. (L.U., Hon. Subsid.)

6. What is meant by the velocity of an ion in electrolysis, and how is it measured?

7. Describe Kohlrausch's method of determining the resistance of an electrolyte; and explain how from a knowledge of the conductivity of salt solutions, the degree of dissociation of a solution of given strength is usually calculated. (L.U., B.Sc. Internal.)

8. Explain how the velocities of the ions in an electrolyte have been ascertained, and describe a method of directly observing an ionic velocity. (L.U., B.Sc. Internal.)

9. Find a relation between the rate of change with temperature of the electromotive force of a reversible cell and the other constants of the cell. (L.U., B.Sc. Hon. Internal.)

10. A sphere of unit radius contains a solution of hydrochloric acid in which the density of the acid is 10^{-4} gramme per cubic centimetre. Calculate the electric force in volts per centimetre at the surface of the sphere if 1 per cent. of the chlorine ions were removed from the solution, the electrochemical equivalent of hydrogen being 0·000104 gramme. (Atomic weight of chlorine=35·5, hydrogen=1·01.) [$1{\cdot}003\times10^{10}$ volts per cm.]

Hence, show that it would be impossible by employing forces of the order of a volt per centimetre to produce any separation of H and Cl ions in the solution that could be estimated chemically. (L.U., B.Sc. Ext.) [$4{\cdot}06\times10^{-16}$ gm. cl.]

11. Two liquid resistances, A and B, of 5 and 10 ohms respectively, are connected in parallel, and a battery of electromotive force 8 volts and 2 ohms internal resistance is used to send a current through them.

Find the currents in the two liquids, being given that the electromotive force of polarisation is 0·1 volt in A and 1·8 volts in B. (L.U., B.Sc. Ext.) [A, 1·03, B, 0·345 amp.]

12. Explain how the electromotive force of a cell may be deduced from the quantities of heat evolved in the chemical reactions that take place in the cell, and show that the correction for the temperature-variation of the electromotive force is $T\frac{dE}{dT}$. (L.U., B.Sc. Hon. Ext.)

13. Give an account of experiments to determine the transport numbers for ions in electrolysis. (L.U., B.Sc. Hon. Internal.)

14. Show how the velocity of electrolyte ions in an electric field can be calculated from measurement of the specific resistance, and of the transport ratio.

Describe, mentioning necessary precautions, experiments by which this velocity is directly measured. (L.U., B.Sc. Internal.)

15. What is meant by the term Solution Pressure used in connection with Voltaic cells? Show how an expression for the electromotive force developed has been deduced through this conception. (L.U., B.Sc. Hon. Internal.)

16. Explain the meaning of the expression "u and v the mobilities of the ions in electrolysis."

Show that if in the electrolysis of a solution $10{\cdot}36\times10^{-6}$ gramme equivalents of each ion are liberated by the passage of an ampere for a second,

$$u+v=10{\cdot}36\times10^{-6}k/N$$

when k is the conductivity of the electrolyte and N the number of gramme equivalents of dissolved salt per cubic centimetre of the solution. (L.U., B.Sc. Ext.)

17. Deduce from thermodynamical principles an expression for the rate of change of the E.M.F. of a reversible cell with temperature.

A certain reversible cell has an E.M.F. of 1 volt at 0° C. Assuming that the energy provided by the chemical reactions occurring in the cell amounts to 0·26 calories per coulomb passing through it, calculate the change in its E.M.F. when its temperature is raised from 0° C. to 1° C. (L.U., B.Sc. Ext.) [0·000324 volt.]

Examples VII

1. What are the Peltier and Thomson coefficients and how are they represented on the thermo-electric diagram? (L.U., B.Sc. Internal.)

2. Give the theory of thermo-electromotive force, and show that the coefficient of the Peltier effect

$$=T\cdot\frac{dE}{dT},$$

T being the absolute temperature of the junction, and E the whole E.M.F. acting in the circuit. (L.U., B.Sc. Hon.)

3. What is the Thomson thermo-electric effect? Give the reasoning which led to its discovery.

4. Write a short essay on the use of thermo-couples to measure temperature, and describe the determination of the temperature at which a molten mixture of two metals solidifies.

5. Prove that the coefficient of the Peltier effect at a given junction is the product of the absolute temperature of the junction and the rate of change of the whole E.M.F. of the circuit with the temperature of that junction. (L.U., B.Sc. Hon. Internal.)

6. What is meant by the specific heat of electricity? Assuming that the E.M.F. of a circuit of two metals with the cold junction kept at constant temperature varies with the temperature of the hot junction according to a parabolic law, show that the difference of the specific heats of electricity in the two metals is proportional to the absolute temperature. (L.U., B.Sc. Hon. Internal.)

7. The thermo-electric power of iron is 17·34 micro-volts per degree at 0° and 12·47 at 100°, that of copper is 1·36 at 0° and 2·31 at 100°. Construct a thermo-electric diagram for these metals, lead being the standard; and state how the amounts of heat absorbed and given out in the different parts of a copper-iron circuit with its junctions at 0° and 100°, when there is a current of 1 ampere, are shown in the diagram.

Calculate also the electromotive force in volts. (L.U., B.Sc. Internal.) [0·1307 volt.]

8. Explain clearly what is meant by the "specific heat of electricity."

Along a metal rod whose area of cross-section is 1 sq. cm. there is a uniform temperature gradient of 1° C. per centimetre. The specific resistance of the material of the rod is 150 microhms per cm. cube. When a current of 0·05 ampere is sent from the hot to the cold end the temperature gradient is unaltered. Calculate the specific heat of electricity for this metal. (L.U., B.Sc. Internal.) [$\sigma = -7{\cdot}5 \times 10^{-6}$ volt per degree centigrade.]

9. The E.M.F. in a simple thermo-electric circuit, one junction of which is heated while the other is kept at 0° C., is given by the expression $E = bt + ct^2$ where t is the temperature of the hot junction. Determine the neutral temperature, and the Peltier and Thomson effects in the circuit.

Explain the theory on which these determinations are made. (L.U., B.Sc. Hon. Internal.) [$t_n = -b/2c$; $\pi = (t+273)(b+2ct)$; $\sigma_a - \sigma_b = 2c(t+273)$.]

10. Describe how you would measure the thermo-electric power of copper with respect to iron at various temperatures.

The thermo-electric power of iron is 17·5 micro-volts per degree at 0° C. and is zero at 360° C.; the thermo-electric power of copper is 5 micro-volts per degree at 450° C., and zero at −50° C. Draw the thermo-electric lines for these elements and deduce a value for the E.M.F. in a copper-iron circuit when the cold junction is at 0° C., and the hot junction is at the neutral temperature. (L.U., B.Sc. Internal.) [2465 microvolts.]

Examples VIII

1. Describe some form of suspended coil galvanometer stating the conditions under which it is (1) dead beat or (2) ballistic.

2. Deduce an expression for the work done in taking a unit magnetic pole round a closed contour embracing an electric current. The internal and external radii of an anchor ring are 9 and 10 cm. respectively. If the ring is wound with 1,000 turns of wire carrying a current of 2 amperes, find the magnetic intensity at a point 9·2 cms. from the axis. (L.U., B.Sc. Ext.) [43·48.]

3. Prove the formula for calculating the magnetic field inside a long helix at points distant from the ends. Suggest a method for measuring the field inside experimentally. (L.U., B.Sc. Internal.)

4. Find an expression for the magnetic potential at any point due to an electric current flowing round a closed circuit. Hence, or otherwise, calculate the galvanometer constant of the Helmholtz form of tangent galvanometer, in which two coils are placed parallel to each other at a distance apart equal to the radius of either. (L.U., B.Sc. Internal.) [$32\pi n/a\sqrt{125}$ abs.]

5. A solenoid of 1000 turns is wound uniformly in a single layer on a tube 50 cm. long and 10 cm. diameter. Determine the strength of the magnetic field in C.G.S. units, (*a*) at the centre of the solenoid, (*b*) at the centres of ends when a current of 0·1 ampere flows through the wire. (L.U., B.Sc. Hon.) [(*a*) 2·494; (*b*) 1·250.]

6. An electric current of 1 ampere flows round a circular circuit, the radius of which is 10 cm. Determine the strength and direction of the magnetic field at a point on the line drawn through the centre of the circle perpendicular to its plane and 10 cm. distant from the plane of the ring. [0·0222 c.g.s.]

7. Explain the method of measuring the earth's magnetic field by means of an inductor and ballistic galvanometer. (L.U., B.Sc. Internal.)

8. Two circular coils of wire are placed with their planes parallel to each other at a distance of 5 cm. apart. The larger coil has a radius of 10 cm. and 30 turns of wire, the smaller a radius of 2 cm. and 20 turns of wire. Calculate approximately in grammes weight the mechanical force between the coils when a current of 1 ampere is passed through both, proving any formula used.

Show how the principles thus illustrated are applied practically in the ampere-balance. (L.U., B.Sc. Hon. Ext.) [0·00415 gm. wt.]

9. Find an expression for the mutual potential energy of a magnetic shell and an external magnetic system. Show that if the shell forms a closed surface it will exert no action on a magnet inside it. (L.U., B.Sc. Hon. Internal.)

10. Find the law of refraction of magnetic lines at a surface at which the permeability of the medium changes.

Draw incident and refracted lines for media whose permeabilities are in the ratio 1·5, when the incident line makes an angle of 45 with the normal to the surface (*a*) in the medium (1), (*b*) in the medium (2). (L.U., B.Sc. Hon. Ext.) [(*a*) 33° 41′; (*b*) 56° 19′.]

EXAMPLES IX

1. In what respects do the magnetic properties of iron and steel differ? Define the terms intensity of magnetisation (I), induction (B), and magnetic force (H).

How do you obtain the relation

$$B = H + 4\pi I,$$

either theoretically or experimentally? (L.U., B.Sc. Ext.)

2. Show in what features a magnet circuit is analogous to an electric circuit. In what respects does the analogy fail? (L.U., B.Sc. Internal.)

3. Show that the work per cubic centimetre performed in taking a specimen of iron through a cycle of magnetisation is represented by the area of the cycle upon the H—I diagram. Describe how the energy loss due to hysteresis may be determined for a given material. (L.U., B.Sc. Hon. Internal.)

4. Define magnetic induction B and magnetising force H, and give an account of an experimental method of determining their relation for a specimen of soft iron. (L.U., B.Sc. Internal.)

5. Explain what is meant by residual magnetism, coercive force, permeability.

Draw a curve showing the manner in which the magnetisation of a soft iron rod varies as the magnetising field is taken through a cycle, and state in a general way how from this diagram you would obtain the residual magnetism and coercive force.

6. Define magnetic force H and magnetic induction B. Show that the energy per unit volume of the magnetic field between two plane poles is given by $\frac{BH}{8\pi}$. (L.U., B.Sc. Ext.)

7. Discuss the effects of and the methods of dealing experimentally with free magnetism in the measurement of magnetic permeability. Find an expression for the effect of a thin radial crevasse upon the magnetisation of an anchor ring. (L.U., B.Sc. Hon. Internal.)

8. What is the general character of the magnetic permeability of iron (*a*) in very strong, (*b*) in very weak fields? How has the latter been experimentally investigated? (L.U., B.Sc. Hon. Internal.)

9. What is meant by hysteresis, and by a cycle of magnetisation?

Prove that the area of the H, B cycle denotes 4π times the energy dissipated per c.c. of metal during each magnetic cycle. (L.U., B.Sc. Internal.)

10. A long solenoid of ten turns to the centimetre contains an iron rod 2 cm. diameter cut in two. Find the force necessary to separate the two halves of the rod when a current of 3 amperes is flowing in the solenoid; given that on reversing this primary circuit 60 micro-coulombs flow through a secondary circuit of ten turns wound round the iron rod, and having a total resistance of 100 ohms. (L.U., B.Sc. Internal.) [$1{\cdot}139\times10^7$ dynes.]

11. Describe the effect of temperature on (1) the magnetisation of iron under small magnetising forces; (2) the maximum intensity of magnetisation of the iron.

12. Describe the ballistic method of determining the relation between the magnetisation and the magnetising field in iron in the form of a ring.

13. Explain some method of measuring the energy lost through magnetic hysteresis, and describe some of the principal results of experiment.

Examples X

1. Define the coefficient of self-induction or inductance of a circuit. Calculate approximately the inductance of a long straight cylindrical solenoid of radius r and length l, wound uniformly with N turns of wire per unit length. [$4\pi^2r^2N^2l\times10^{-9}$ henries.]

2. Discuss the production of electric oscillations in a circuit containing capacity and inductance.

3. Describe some accurate method for the measurement of a high resistance.

A condenser of capacity 0·5 microfarad and resistance 10 megohms is charged to a certain potential and then insulated. Find the time the potential will take to fall to half its original value ($\log_e 2=0{\cdot}6931$). (L.U., B.Sc. Internal.) [3·465 sec.]

4. Define "self-inductance of a circuit," and describe in detail any two phenomena which depend upon it. (L.U., B.Sc. Internal.)

5. Establish the equation representing the discharge of a condenser through an inductive circuit. Under what conditions will the discharge be oscillatory? Determine the frequency of the oscillation in the case of a Leyden jar of capacity 0·001 microfarad, discharged by "tongs," the circuit having an inductance of 0·003 henry. (L.U., B.Sc. Hon.) [91910 cycles per sec.]

6. Assuming the equation

$$Ri=E-nS\frac{dB}{dt}$$

for the current i in a solenoid of n turns wound round a long iron bar of section S, show that $B=H+4\pi I$, where I is the intensity of magnetisation of the iron and H the magnetic force due to the current in the solenoid. (L.U., Ext. Hon.)

7. A solenoid coil 70 cm. in length, wound with 30 turns of wire per centimetre, has a radius of 4·5 cm. A second coil of 750 turns is wound upon the middle part of the solenoid. Calculate the coefficient of self-induction of the solenoid and the coefficient of mutual induction of the two coils. Will the inductance of the solenoid be affected by short circuiting the ends of the secondary coil? (L.U., B.Sc. Internal.) [50·36 m.h.; 17·99 m.h.; yes.]

8. Find whether the discharge of a condenser through an inductive circuit is oscillatory when

(*a*) Capacity = 2 microfarad, L = 0·15 henry, and R = 150 ohms. [Yes; 279·5]
(*b*) C = 1·5 m.f., L = 0·015 hy., and R = 1000 ohms. [No]
(*c*) C = 10^{-6} m.f., L = 0·0125 hy., and R = 100 ohms. [Yes; $1{\cdot}423\times10^6$]

and when oscillatory, find the frequency.

9. Find approximately the frequency of oscillation when a condenser of 0·75 microfarad capacity discharges through a circuit consisting of a solenoid of 1200 turns and length 80 cm. wound upon a long iron rod of diameter 0·75 cm. and permeability 850. [630·6 cycles per sec.]

10. State the law of electromagnetic induction and explain how it is mathematically formulated.

An earth inductor (200 turns of wire of mean diameter 40 cm.) is placed with the diametral axis about which it revolves horizontal and in the magnetic meridian. If the value of H is 0·2 gauss and the angle of dip is 60° find an expression for the voltage generated in the coil when it rotates 10 times per second. (L.U., B.Sc. Ext.) [$V=54{\cdot}7\sin 20\pi t$ m.v.]

EXAMPLES XI

1. Two points, to which an alternating E.M.F. is applied, are connected by a circuit containing capacity and inductance: give a method by which the current and its lag behind the E.M.F. may be determined.

Find the current if maximum E.M.F.$=200$ volts, frequency$=50$, resistance $=10$ ohms, inductance$=0{\cdot}1$ henry, capacity$=1$ microfarad. [$0{\cdot}0897 \sin (100\pi t + \pi/2)$.]

2. Distinguish between the mean value and the root mean square value of an alternating current. Find the relation between them.

Prove that the power absorbed by a coil traversed by an alternating current is EI $\cos \theta$, where E and I are the root mean square values of the E.M.F. and current respectively, and θ is the difference in phase between these two quantities.

3. Find an expression for the current at any moment in a circuit of given resistance and capacity when subject to a simple harmonic E.M.F.

4. An alternating E.M.F. is applied to a circuit A. Show that the presence of a neighbouring circuit B has the apparent effect of increasing the resistance of A and diminishing its self-induction. (L.U., B.Sc. Hon.)

5. Describe Kelvin's ampere-balance, and explain its advantages as a standard for calibrating ammeters and voltmeters for alternating currents. (L.U., B.Sc. Hon. Ext.)

6. Explain the apparent increase in resistance of a wire with the frequency for a rapidly alternating current. (L.U., B.Sc. Hon. Internal.)

7. Describe the construction of an electrostatic voltmeter. An electrostatic voltmeter gives deflections of 15, 18, and 21 scale divisions for constant potentials of 50, 60, and 70 volts respectively. What deflections will be produced by an alternating electromotive force E $\sin pt$ (*a*) when the amplitude E is 70 volts, and (*b*) when E is 90 volts? (L.U., B.Sc. Ext.) [14·85; 19·09.]

8. Show that two alternating magnetic fields at right angles to each other may be equivalent to a rotating field, and explain how this has been utilised in the construction of electric motors.

9. What is meant by *impedance*? Show how to calculate the current which passes through a circuit of known resistance and inductance.

10. What is an alternating current? Describe some type of instrument suitable for measuring it.

An incandescent lamp is joined in series with a 4 microfarad condenser to an alternating supply of 50 cycles per sec. If the current is ¼ ampere and the potential difference across the lamp terminals is 100 volts, what is the supply voltage? (L.U., B.Sc. Internal.) [222·7 volts (R.M.S.).]

11. Discuss the difficulties encountered in making an accurate determination of the power given to an inductive circuit such as the primary of a transformer, and describe some satisfactory method of making the measurement. (L.U., B.Sc. Hon. Ext.)

12. An alternating E.M.F. of 200 volts and 50 periods per second is applied to a condenser in series with a 20 volt 5 watt metal filament lamp. Find the capacity of the condenser required to run the lamp. (L.U., B.Sc. Hon. Internal.) [4 microfarads.]

13. Show how the closing of the secondary circuit affects the phase and current in the primary of an ironless transformer due to an applied alternating electromotive force. (L.U., B.Sc. Hon. Internal.)

14. An alternating electromotive force is applied to a non-inductive circuit consisting of a resistance and a capacity in series. Find an expression for the current finally produced.

A condenser of capacity 1 mfd. and a resistance of 50 ohms are connected in series with an alternating supply of 100 volts amplitude and frequency 50 cycles per second. The addition to the circuit of an inductive coil of negligible resistance is found to cause the current to increase. Explain this and find the value of the added self-inductance to give maximum current. State also the value of this maximum current. (L.U., B.Sc. Internal.) [10·13 henries; 2 amp.]

15. Explain the method of solving alternating-current problems with the aid of complex quantities. Investigate by such a method the behaviour of *either* (*a*) a parallel resonance circuit (consisting of a capacity C in parallel with a coil of resistance R and self-inductance L), *or* (*b*) an alternating-current bridge suitable for the measurement of self-inductance. (L.U., B.Sc. Special Internal.)

Examples XII

1. Describe a method for determining the ohm in absolute measure.

2. What is meant by the expression "the dimensions of a physical quantity"? Taking as your fundamental quantities time, length, and force, deduce the dimensions of energy and electrostatic potential. (L.U., B.Sc. Ext.)

3. Find the dimensions in mass, length and time, of electric current, capacity, and inductance, on both the electrostatic and the electromagnetic systems. What important relation involving μ and k can be deduced from these expressions? (L.U., B.Sc. Hon. Subsid.)

4. Determine the dimensions of electric resistance in electrostatic and in electromagnetic units.

5. Give Lorenz's method of determining the value of the ohm in absolute measure.

6. Examine the dimensions of the quantity $k\mu$ where μ denotes permeability and k dielectric constant.

7. Explain why the ratio of the electromagnetic units to the electrostatic units is so intimately connected with the velocity of light.

8. Describe some one method of determining experimentally the quantity "v" involved in the ratio of the two systems of electric units, explaining the precautions necessary for an accurate result. (L.U., B.Sc. Hon. Ext.)

9. Describe some methods by which the units of current and electromotive force may be found in the electromagnetic system. (L.U., B.Sc. Ext.)

10. According to the usual definitions, the dimensions of capacity on the electromagnetic system are those of the reciprocal of an acceleration, while on the electrostatic system they are simply a length. Show how these results are obtained, and explain the apparent discrepancy. (L.U., B.Sc. Internal.)

11. Define the terms—magnetomotive force, magnetic flux, and reluctance of a magnetic circuit. Find the dimensions of these quantities. (L.U., B.Sc. Hon. Internal.)

12. Derive the dimensions of electric charge (*a*) in terms of those of dielectric constant, mass, length and time, and (*b*) in terms of those of permeability, mass and length. Equate the two expressions and discuss the significance of the result. (L.U., B.Sc. Subsid. Internal.)

Examples XIII

1. How has it been shown experimentally that the current in discharging a condenser may be alternating in character?

2. Give a short account of how the wave-length of electromagnetic waves in air has been determined. (L.U., B.Sc. Internal.)

3. What method would you adopt to determine the dielectric constant of a slightly conducting liquid? (L.U., B.Sc. Hon. Internal.)

4. Explain the conditions necessary for periodic electromagnetic disturbances to be transmitted through a cable without distortion. (L.U., B.Sc. Hon.)

5. Prove that in a plane electromagnetic wave the electric field and the magnetic field are in the wave front and in directions at right angles to one another.

6. Show that the discharge of a condenser may be oscillatory, and describe a method by which the period of the oscillations may be measured.

7. Describe experimental methods of finding the velocity of electromagnetic waves along wires. (L.U., B.Sc. Hon. Ext.)

8. Write down the fundamental equations of the electromagnetic field, and deduce from them the velocity of propagation of a plane electromagnetic wave in a medium of dielectric constant k and permeability μ. (L.U., B.Sc. Hon. Internal.)

9. Describe the production of electromagnetic waves by means of Hertz's oscillators, and show that there is necessarily a relative change in phase of the electric and magnetic components accompanying the progress of the disturbance outward from the oscillator. (L.U., B.Sc. Hon. Internal.)

10. Describe any three methods of detecting electric oscillations, and discuss their relative advantages. (L.U., B.Sc. Hon. Internal.)

11. Show that the magnetic effects of a current may be regarded as due to the motion of the Faraday tubes, and find the paths along which the energy travels. (L.U., B.Sc. Hon. Internal.)

Examples XIV

1. Describe experiments which show that the cathode rays are small particles of negative electricity.
2. Describe a method of measuring the velocity of the ions in gases.
3. Give a brief account of the principal phenomena of cathode rays, and explain how the charge carried has been determined.
4. Give an account of some recent determination of e the electronic charge, and give the theory of the method.
Taking e as $4 \cdot 7 \times 10^{-10}$ *electrostatic* units, the electrochemical equivalent of hydrogen as 0·0000104 gm. per coulomb, and the standard density of hydrogen as 0·00009 gm. per c.c., estimate the number of molecules per c.c. of hydrogen under standard temperature and pressure. (L.U., B.Sc. Hon.) [$2 \cdot 762 \times 10^{19}$.]
5. How may the electrical conductivity of an ionised gas be determined? What is meant by the saturation current? (L.U., B.Sc. Ext.)
6. Give two methods of finding the ratio of electric charge to mass in the case of an ion in a gas. (L.U., B.Sc. Hon. Internal.)
7. Find the conditions that the effect of a charge situated upon a small spherical drop shall be equal and opposite to that due to surface tension. (L.U., B.Sc. Hon. Internal.)
8. Describe how the elementary ionic charge has been measured.
A charged oil drop is suspended by a uniform field of 300 volts per cm., so that it neither rises nor falls. Find the charge on it, assuming its mass to be $9 \cdot 75 \times 10^{-12}$ gram. (L.U., B.Sc. Ext.) [$9 \cdot 565 \times 10^{-9}$ e.s.u.]
9. A small mass m moves with speed v into (*a*) a magnetic field, (*b*) an electric field, each field being perpendicular to the initial direction of motion of the charge. Find the path of the charged particle in each case and explain how, by arranging the fields in a suitable manner, the quotient e/m may be found.
How have such methods been applied to the resolution of isotopes? (L.U., B.Sc. Hon.) $\left[\text{Circular, } r=\frac{mv}{He}\text{; parabolic, semi } l.r.=\frac{mv^2}{Ze}.\right]$
10. Describe how the ratio of the charge to the mass of an electron has been determined.
Electrons pass from a hot cathode across a vacuous space to a collecting plate which is maintained at a potential of 10 volts relative to the cathode. Find the velocity of the electrons when they reach the plate, assuming that e/m for electrons is $1 \cdot 77 \times 10^7$ e.m.u./gm. (L.U., B.Sc. Internal.) [1·881 cm. sec.$^{-1}$.]
11. Give several well-established examples of isotopes, and indicate the physical means by which they were discovered. (L.U., B.Sc. Subsid.)
12. Explain the diffraction of X-rays by a crystal. Examine especially the case when parallel atom planes are of alternating strengths and spacings, and give examples. (L.U., B.Sc. Hon.)
13. Explain the method by which an absolute determination of an X-ray wave-length has been made. (L.U., B.Sc. Hon.)
14. Show that the electrification of a sphere of water is favourable to the condensation of water vapour upon it. To what potential should a drop of water of 1 mm. radius be raised that its vapour pressure may be the same as that at a flat unelectrified surface, given that the surface tension of water =75 dynes per cm. (L.U., B.Sc. Hon. Internal.) [19·42 e.s.u.]
15. Describe the mass spectrograph of Aston, and explain its special advantages in the detection of isotopes.
Give the theory of the focussing property of the instrument. (L.U., Special Int.)

Examples XV

1. Write an essay on radium, in particular discussing the changes which this substance undergoes.
2. Describe experiments by which the nature of the rays emitted by radium has been determined.

3. Give an account of the various kinds of radiation emitted by a solid compound of radium, explaining how the properties of the different rays may be investigated experimentally. (L.U., B.Sc. Hon. Ext.)

4. Give an outline of the theory of the disintegration of radioactive materials, and deduce equations showing the amounts of two consecutive products present at any time subsequent to the isolation of the higher product. (L.U., B.Sc. Hon. Internal.)

5. Find by calculation the expression for the growth of a radioactive substance (*a*) when the parent substance undergoes very slow transformation, and (*b*) when the parent substance is one whose decay is rapid. (L.U., B.Sc. Hon. Subsid.)

6. Describe experiments by which the number and the total charge of the α particles emitted per second by a known quantity of a radioactive substance have been determined, and mention briefly some of the more important deductions which may be made from the results of these experiments. (L.U., Special Int.)

Examples XVI

1. Describe the application of the electron theory to the explanation of electric conductivity.

2. When radiation of frequency n falls on matter electrons are given out possessing a maximum velocity given by the formula :

$$\tfrac{1}{2}mv^2 = h(n - n_0)$$

Explain how this has been verified.

Radiation falls on a target within a solenoid of 20 turns per cm., carrying a current of $2\tfrac{1}{2}$ amperes. The electrons then move in circles, the largest of which are of radius 1 cm. Find the frequency of the radiation assuming the following data :—

$e/m = 1{\cdot}77 \times 10^7$ em. units. $\qquad h = 6{\cdot}55 \times 10^{-27}$ ergs.-sec.

$m = 0{\cdot}9 \times 10^{-27}$ gm. $\qquad n_0 = 0$.

(L.U., B.Sc. Hon.) [$8{\cdot}497 \times 10^{16}$ cycles per sec.]

3. Find the magnetic field due to a point charge moving with uniform velocity small compared with that of light. [$(ev \sin\theta)/r^2$.]

4. Give a short account of the Zeeman effect, and of its explanation on the simple electronic hypothesis.

5. Give an account of Bohr's theory of atoms, considering more especially the hydrogen spectra. (L.U., B.Sc. Hon.)

6. Distinguish between diamagnetic, paramagnetic and ferromagnetic substances stating briefly your reasons.

7. Give a theory of the production of Röntgen rays, and discuss how it explains the leading experimental results. (L.U., B.Sc. Hon. Ext.)

8. Give an account of the Stern-Gerlach experiment on the magnetic deviation of atomic rays and discuss its theoretical significance. (L.U., B.Sc. Special Ext.)

9. Discuss the nature of " electrical mass," and describe experiments in which the variation of electronic mass with velocity has been accurately investigated. (L.U., B.Sc. Special Internal.)

10. What do you understand by the terms " Bohr magneton " and " Weiss magneton " ? Give an account of experimental work on the measurement of the magnetic moments of atoms. (L.U., B.Sc. Special Ext.)

INDEX